Dynamic brain networks explored by structure-revealing methods

THÈSE N^{O} 6274 (2014)

PRÉSENTÉE LE 27 JUIN 2014

À LA FACULTÉ DES SCIENCES ET TECHNIQUES DE L'INGÉNIEUR

GROUPE VAN DE VILLE

PROGRAMME DOCTORAL EN BIOTECHNOLOGIE ET GÉNIE BIOLOGIQUE

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

POUR L'OBTENTION DU GRADE DE DOCTEUR ÈS SCIENCES

PAR

Nora LEONARDI

acceptée sur proposition du jury:

Prof. N. Stergiopulos, président du jury
Prof. D. Van De Ville, directeur de thèse
Dr S. Achard, rapporteuse
Prof. M. Greicius, rapporteur
Prof. J. D. R. Millán Ruiz, rapporteur

ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

Suisse
2014

Abstract

The human brain is a complex system able to continuously adapt. How and where brain activity is modulated by behavior can be studied with functional magnetic resonance imaging (fMRI), a non-invasive neuroimaging technique with excellent spatial resolution and whole-brain coverage. FMRI scans of healthy adults completing a variety of behavioral tasks have greatly contributed to our understanding of the functional role of individual brain regions. However, by statistically analyzing each region independently, these studies ignore that brain regions act in concert rather than in unison. Thus, many studies since have instead examined how brain regions interact. Surprisingly, structured interactions between distinct brain regions not only occur during behavioral tasks but also while a subject rests quietly in the MRI scanner. Multiple groups of regions interact very strongly with each other and not only do these groups bear a striking resemblance to the sets of regions co-activated in tasks, but many of these interactions are also progressively disrupted in neurological diseases. This suggests that spontaneous fluctuations in activity can provide novel insights into fundamental organizing principles of the human brain in health and disease. Many techniques to date have segregated regions into spatially distinct networks, which ignores that any brain region can take part in multiple networks across time. A more natural view is to estimate dynamic brain networks that allow flexible functional interactions (or connectivity) over time. The estimation and analysis of such dynamic functional interactions is the subject of this dissertation. We take the perspective that dynamic brain networks evolve in a low-dimensional space and can be described by a small number of characteristic spatiotemporal patterns. Our proposed approaches are based on well-established statistical methods, such as principal component analysis (PCA), sparse matrix decompositions, temporal clustering, as well as a multiscale analysis by novel graph wavelet designs. We adapt and extend these methods to the analysis of dynamic brain networks. We show that PCA and its higher-order equivalent can identify co-varying functional interactions, which reveal disturbed dynamic properties in multiple sclerosis and which are related to the timing of stimuli for task studies, respectively. Further we show that sparse matrix decompositions provide a valid alternative approach to PCA and improve interpretability of the identified patterns. Finally, assuming an even simpler low-dimensional space and the exclusive temporal expression of individual patterns, we show that specific transient interactions of the medial prefrontal cortex are disturbed in aging and relate to impaired memory.

Keywords: brain imaging, fMRI, resting state, functional connectivity, dynamic networks, graph wavelets, matrix factorization, tensor decomposition, sparsity, multiple sclerosis, aging

Résumé

Le cerveau humain est un système complexe capable de s'adapter en permanence. Comment et où l'activité cérébrale est modulée par le comportement peut être étudié avec l'imagerie par résonance magnétique fonctionnelle (IRMf), une technique de neuro-imagerie non invasive caractérisrée par une excellente résolution spatiale et qui permet d'explorer le cerveau en entier. Des études sur des adultes sains exécutant une variété de tâches comportementales ont beaucoup contribué à notre compréhension de la fonction des différentes régions du cerveau. Cependant, en analysant chaque région indépendamment, ces études négligent le fait que les régions du cerveau agissent collectivement plutôt qu'individuellement. Aussi, de nombreuses études ont récemment examiné la façon dont les régions du cerveau interagissent. Il s'avére que, des interactions structurées entre régions distinctes du cerveau ne sont pas seulement évidentes quand un sujet est stimulé par une tâche comportementales, mais aussi quand le sujet est au repos. Des sous-groupes de régions distinctes interagissent très fortement les unes avec les autres, et ces groupes ont non seulement une ressemblance frappante avec les groupes de régions co-activés par des tâches, mais beaucoup de ces interactions sont également progressivement perturbées dans des maladies neurologiques. Ceci suggère que les fluctuations spontanées de l'activité cérébrale peuvent fournir de nouveaux éclairages sur les principes d'organisation du cerveau sain et pathologique. A ce jour, la plupart des méthodes séparént les régions dans des réseaux spatialement distincts, mais ignorent le fait que toute région du cerveau peut faire partie de plusieurs réseaux àu cours du temps. Une approche plus générale consiste à estimer des réseaux cérébraux dynamiques qui permettent d'établir une connectivité fonctionnelle évolutive dans le temps. L'estimation et l'analyse de cette connectivité fonctionnelle dynamique sont le sujet de cette thèse. Nous partons du point de vue que les réseaux dynamiques du cerveau évoluent dans un espace de petite dimension et peuvent être décrits par un petit nombre de modèles spatio-temporelles caractéristiques. Nous proposons des approches basées sur des méthodes statistiques établies, telles que l'analyse en composantes principales (ACP), les décompositions parcimonieuses, les méthodes de regroupement dans le temps et une analyse multi-échelles originale d'ondelettes sur graphes. Nous adaptons ces méthodes à l'analyse des réseaux cérébraux dynamiques. Nous montrons que l'ACP et son analogue d'ordre supérieur peuvent identifier des connexions fonctionnelles co-variantes, qui révèlent des propriétés dynamiques perturbées dans la sclérose en plaques et qui sont liées à l'occurrence des stimulis dans une étude de tâche comportementale, respectivement. De plus, nous montrons que les décompositions parcimonieuses offrent une alternative valable à l'ACP et améliorent l'interprétation des modèles identifiés. Enfin, sous l'hypothèse d'un espace de petite dimension encore plus simple et de la contribution temporelle unique d'un seul modèle, nous montrons que des interactions transitoires spécifiques du cortex préfrontal médial sont perturbées dans le vieillissement et sont liées à la performance mnésique.

Mots clés : imagerie cérébrale, IRMf, état de repos, connectivité fonctionnelle, réseaux dynamiques, ondelettes sur graphhes, factorisation de matrices, factorisation de tenseurs, parcimonie, sclérose en plaques, vieillissement

Zusammenfassung

Das menschliche Gehirn ist ein komplexes System, das sich kontinuierlich anpasst. Wie und wo sich die Aktivität des Gehirns durch menschliches Verhalten ändert, kann mit der funktionellen Magnetresonanztomographie (fMRT) untersucht werden, einem nichtinvasiven bildgebenden Verfahren mit hervorragender räumlicher Auflösung, das die Aktivität im ganzen Gehirn misst. FMRT-Studien mit gesunden Erwachsenen, die eine Vielzahl von Wahrnehmungs- und Verhaltensaufgaben ausführten, haben wesentlich zu unserem Verständnis der Funktion der einzelnen Hirnareale beigetragen. Die unabhängige statistische Analyse der Aktivität der einzelnen Hirnareale ignoriert jedoch, dass Hirnareale zusammen und nicht unabhängig voneinander arbeiten. Daher haben viele Studien in jüngerer Zeit untersucht, wie Hirnareale funktionell interagieren. Überraschenderweise treten strukturierte Arealinteraktionen nicht nur während Verhaltensaufgaben auf, sondern auch während die Versuchsteilnehmer ruhig im MRT-Scanner liegen. Mehrere Gruppen von Regionen interagieren besonders stark miteinander und diese Gruppen haben nicht nur eine auffallende Ähnlichkeit zu den Hirnarealen, die auch gemeinsam in einer Verhaltensaufgabe aktiv sind, sondern sie sind auch in neurologischen Krankheiten zunehmend gestört. Dies deutet darauf hin, dass die Analyse spontaner Aktivität neue Einblicke in die Organisationsprinzipien des menschlichen Gehirns in Gesundheit und Krankheit bieten kann. Viele Methoden teilen Hirnareale in räumlich getrennte Netzwerke auf, was ausschliesst, dass eine Gehirnregion über die Zeit Teil mehrerer Netzwerke ist. Eine natürlichere Betrachtungsweise ist jene der dynamischen Netzwerke, in welchen flexible Interaktionen erlaubt sind. Die Analyse solcher dynamischen Interaktionen ist das Thema dieser Dissertation. Aus unserer Sicht entwickeln sich dynamische Gehirn-Netzwerke in einem niedrig-dimensionalen Raum und können durch eine kleine Anzahl charakteristischer Raum-Zeit-Mustern beschrieben werden Unsere vorgeschlagenen Ansätze basieren auf etablierten statistischen Methoden, wie die Hauptkomponentenanalyse (PCA), “sparse“-Matrix Zerlegung, zeitliche Clusteranalyse sowie eine Multiskalenanalyse für Netzwerke mit einem neuen “Wavelet“-basierten Design. Wir passen diese Methoden an und erweitern sie für die Analyse von dynamischen Gehirn-Netzwerken. Wir zeigen, dass PCA und sein Analog für Matrizen höherer Stufen ähnlich fluktuierende Interaktionen identifizieren können, deren dynamischen Eigenschaften in der Multiplen Sklerose gestört sind, bzw. mit den Zeitpunkten der Verhaltensaufgaben zusammenhängen. Weiter zeigen wir, dass “sparseMatrix Zerlegungen eine gute Alternativen zu PCA bieten und die Interpretierbarkeit der identifizierten Muster verbessern. Schliesslich, unter der Annahme eines noch einfacheren niedrig-dimensionalen Raumes und der exklusiven zeitlichen Erscheinung eines einzelnen Musters, zeigen wir, dass bestimmte dynamische Wechselwirkungen des medialen präfrontalen Kortexes im hohen Alter gestört sind und mit verminderter Gedächtnisleistung zusammenhängen.

Stichworte: Gehirnbildgebung, fMRT, Ruhezustand, funktionelle Konnektivität, dynamische Netzwerke, Netzwerk-Wavelets, Matrix-Zerlegung, Tensor-Zerlegung, Sparsity, Multiple Sklerose, Alterung

Success comes from knowing that you did your best
to become your best.

— John Wooden

Acknowledgements

Tout d'abord je tiens à remercier Dimitri, mon directeur de thèse. J'ai appris énormément de toi – de tes connaissances approfondies, de ton désir de creuser plus loin, de ta capacité de voir la situation dans l'ensemble. Merci d'avoir été si impliqué dans mes projets de recherche, et, en même temps, de m'avoir laissé autant de liberté. Merci de ne pas m'avoir enfermé dans le bureau et de m'avoir donné l'opportunité de présenter ma recherche à de nombreuses conférences internationales ainsi que de partir à l'etranger pour un séjour de recherche. J'ai l'impression de partir avec un sac à dos remplis d'outils et d'astuces. Merci pour tout, Dimitri !

I would like to thank Mike, whose lab I integrated for six months during my PhD. Thank you for giving me this opportunity. It has been very enriching to work with you and a great learning experience. I greatly appreciated you challenging me to be ever more succinct and repeatedly to answer the question "so what?". Merci, Mike!

Thanks to all my jury members, Sophie, Mike, Prof. José del R. Millán and Nikos, for accepting this task. Thank you all for having taken the time for my defense, as well as for your helpful comments and suggestions. Sophie, j'ai eu beaucoup de plaisir à interagir avec toi à plusieurs reprises pendant mon travail de thèse.

Merci également à Jonas. Pour tout ce que tu m'as appris, pour ton enthousiasme contagieux et pour ta disponibilité, en Suisse ainsi qu'aux US.

Thanks also to all other former or current members of the MIP lab, with whom I had the pleasure to interact. Djano for taking me skiing. Işik for being a great office mate, conference travel companion and helping me out in my repeated laptop disasters. Jeff for being another great office mate, never getting tired of answering my English questions and for showing us "his" San Francisco. And Zafer, Giulia, Rotem, Yury, Frank, Djalel, Elena M., Kirsten and Naghmeh for creating such a friendly lab environment. Thanks also to Elena N. and Hamid, two students I had the pleasure of working with.

Thanks also to all members of the FIND lab. In particular, thanks to André and Bernard for enriching scientific and non-scientific discussions, thanks to Will for getting me started in the lab and being an inspiration with your devotion, and thanks to Martin and Anna for all our great conversations.

Merci à tous les collaborateurs du CIBM à Genève, particulièrement François. Vielen Dank unserer tollen Sekretärin Ruth. Thanks to our floor and coffee machine colleagues from the BIG lab, especially Daniel and Bek. Thanks to all scientists I interacted with at conferences and who shared their ideas and comments on my research with me. Thanks to Matthias, Helen and Claudia for giving me the

opportunity to be part of their labs in earlier years and helping me develop my scientific skills.

Thanks to the awesome crowd of EPFL Toastmasters, who taught me loads about public speaking and evaluating and also entertained me well during our weekly meetings. Special thanks to Carina, who has been a great source of inspiration and for proof reading parts of my thesis.

Thanks to Xchange and the wider ESN network for enriching my life on campus.

Thanks to the Stanford University Triathlon Team with whom I was fortunate enough to train. Most of all thanks to their coach, Gina, for encouraging me and teaching me much about mental strength.

I am also grateful to the Swiss National Science Foundation (SNSF), the CIBM and the Fondation Leenards and Louis-Jeantet for financially supporting this research. I am especially grateful to the Fonds Jean Falk-Vairant, the A3SFIT and the SNSF without whose financial support my research stay at Stanford University would not have been possible.

Special thanks to all my friends who supported and encouraged me over the past four years and for all the great times we have spent together.

Und nicht zuletzt, ein grosses Dankeschön an meine Eltern und meine Schwester – meine grössten Unterstützer und gleichzeitig ärgsten Kritiker. Ich schätze mich glücklich zu wissen, dass ich immer auf euch zählen darf und dass ihr mich auf meinem Weg (und Umwegen) stets unterstützt.

Contents

List of notations

Abbreviations

ACC	Anterior cingulate cortex
AG	Angular gyrus
aIns	Anterior insula
BOLD	Blood oxygenation level-dependent
CAP	Co-activation pattern
CSF	Cerebrospinal fluid
(d)FC	(Dynamic) functional connectivity
DMN	Default mode network
EDSS	Expanded Disability Status Scale
EEG	Electroencephalography
(f)MRI	(Functional) magnetic resonance imaging
FOV	Field of view
FWHM	Full width half maximum
GM	Gray matter
GMM	Gaussian Mixture Model
GLM	General linear model
HC	Healthy control
HF	Hippocampal formation
HMM	Hidden Markov Model
HOSVD	Higher order singular value decomposition
ICA	Independent component analysis
(I/M/S)FG	(Inferior/middle/superior) frontal gyrus
mPFC	Medial prefrontal cortex
MVPA	Multivariate pattern analysis
MTL	Medial temporal lobe
PCA	Principal component analysis
PCC	Posterior cingulate cortex
PET	Positron emission topography
PHC	Parahippocampal cortex
ROI	Region of interest

RRMS	Relapse-remitting multiple sclerosis
RSN	Resting-state network
SD	Standard deviation
SGWT	Spectral graph wavelet transform
SVD	Singular value decomposition
TE	Echo time
TR	Repetition time
WM	White matter

Mathematical formulations

x	Scalar value
$\mathbf{x} \in \mathbb{R}^I$	Vector of I real values
$\mathbf{X} \in \mathbb{R}^{I \times J}$	Real matrix with I rows and J columns
$\mathcal{X} \in \mathbb{R}^{I \times J \times K}$	Real tensor with I rows, J columns and K slices
$\lvert x \rvert$	Absolute value of real scalar x, modulus of complex scalar x
x_i	ith value of vector $\mathbf{x}$
$\mathbf{x}^T$	Transpose of vector $\mathbf{x}$
$\lvert\lvert\mathbf{x}\rvert\rvert_0$	ℓ_0-norm of a vector, defined as the number of nonzero elements
$\lvert\lvert\mathbf{x}\rvert\rvert_1$	ℓ_1-norm of a vector, $\lvert\lvert\mathbf{x}\rvert\rvert_1 = \sum_i \lvert x_i \rvert$
$\lvert\lvert\mathbf{x}\rvert\rvert_2$	ℓ_2-norm of a vector, $\lvert\lvert\mathbf{x}\rvert\rvert_2 = \sqrt{\sum_i x_i^2}$
$\lvert\lvert\mathbf{X}\rvert\rvert_F$	Frobenius norm of a matrix (matrix ℓ_2-norm), $\lvert\lvert\mathbf{X}\rvert\rvert_F = \sqrt{\sum_{ij} x_{ij}^2}$
$\mathbf{x} \otimes \mathbf{y}$	Outer product of two vectors $\mathbf{x} \in \mathbb{R}^I$ and $\mathbf{y} \in \mathbb{R}^J$, defined as $\mathbf{x} \otimes \mathbf{y} = \mathbf{x}\mathbf{y}^T = \begin{bmatrix} x_1 y_1 & x_1 y_2 & \dots & x_1 y_J \\ x_2 y_1 & x_2 y_2 & \dots & x_2 y_J \\ \vdots & \vdots & \ddots & \vdots \\ x_I y_1 & x_I y_2 & \dots & x_I y_J, \end{bmatrix}$
$\mathbf{x}_j$	Column j of matrix $\mathbf{X}$
$\mathbf{x}_{:jk}, \mathbf{x}_{i:j}, \mathbf{x}_{ij:}$	Column, row and tube vectors of tensor $\mathcal{X}$, see Fig. 1a
$\mathbf{X}_{i::}, \mathbf{X}_{:j:}, \mathbf{X}_{::k}$	Horizontal, lateral and frontal slices of tensor $\mathcal{X}$, see Fig. 1b
$\mathbf{X}_{(n)}$	Matrix unfolding of tensor $\mathcal{X}$ along mode (or dimension) n, see Fig. 1b
$\mathcal{X} \times_n \mathbf{Y}$	Multiplication of tensor $\mathcal{X}$ with matrix $\mathbf{Y}$ along mode n. $\mathcal{Z} = \mathcal{X} \times_n \mathbf{Y}$ can be calculated by "refolding" $\mathbf{Z}_{(n)} = \mathbf{Y}\mathbf{X}_{(n)}$.
$\mathbf{I}$	Identity matrix, a square matrix with ones on the diagonal and zeros elsewhere
$x \sim \mathcal{N}(\mu, \sigma)$	x follows a normal distribution with mean μ and standard deviation σ
$\mathbf{x} \sim \mathcal{N}(\mu, \Sigma)$	$\mathbf{x} \in \mathbb{R}^I$ follows a multivariate normal distribution with mean vector $\mu \in \mathbb{R}^I$ and covariance matrix $\Sigma \in \mathbb{R}^{I \times I}$

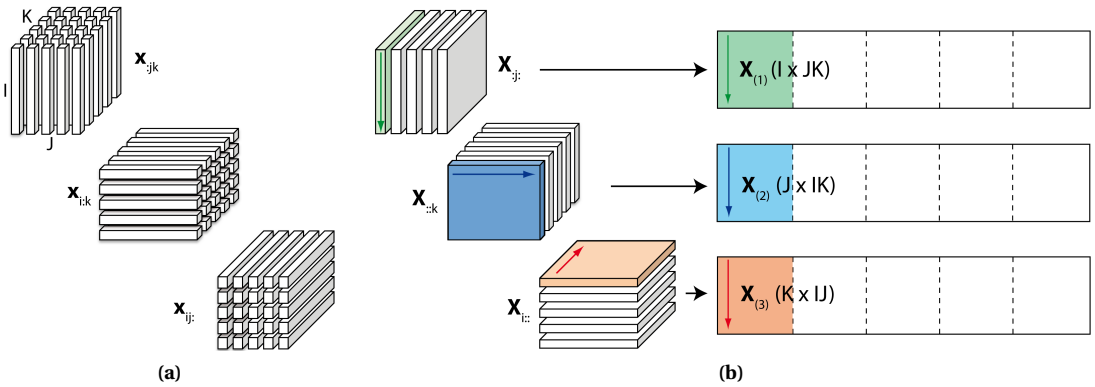


Figure 1 – (a) Column, row, and tube vectors of a third-order tensor $\mathcal{X} \in \mathbb{R}^{I \times J \times K}$ for the case $I = J = K = 5$; (b) lateral, frontal and horizontal slices, and schematics of how the matrix unfoldings $\mathbf{X}_{(1)}$, $\mathbf{X}_{(2)}$, and $\mathbf{X}_{(3)}$ are obtained.

1 Introduction

1.1 Motivation

BRAIN IMAGING METHODS produce rich and complex data on the structure and dynamics of human brain networks. While structural brain connectivity describes the physical "wiring" of the brain and its communication paths, functional connectivity (FC) describes how brain regions interact. FC is estimated by calculating the statistical dependency between the activity of two brain regions over time, such as their temporal correlation, and brain networks can then be defined by combining all pair-wise structural links or functional interactions. Analyses of such networks have provided many novel insights into the structural and functional organization of the human brain. In this dissertation we focus on functional brain networks estimated from functional magnetic resonance imaging (fMRI), a non-invasive neuroimaging technique with excellent spatial resolution and whole-brain coverage.

While FC does not indicate neuronal communication, candidate neuronal underpinnings exist (Nir *et al.*, 2007; He *et al.*, 2008; Britz *et al.*, 2010) and the underlying structural network is thought to shape, but not limit, functional networks (Honey *et al.*, 2007; Greicius *et al.*, 2009; Sporns, 2013). The relationship between structural and functional networks is, however, indirect, which suggests that FC reflects a mix of structurally-constrained and state-dependent interactions (Buckner *et al.*, 2013).

Traditionally, a single functional network has been estimated by correlating brain activity across the entire duration of the scan (we will refer to this type of estimation as "average"), but recent evidence demonstrates the flexible and dynamic nature of these networks. Variability in brain activity and connectivity across time has been explored in electrophysiological recordings already earlier on (e.g. Arieli *et al.*, 1996; Makeig *et al.*, 2004), but, because of its lower temporal resolution, has only more recently been studied using fMRI. These studies have shown that both strength and topology of functional brain networks vary either in response to changes in the external environment or spontaneously while a subject is resting inside the scanner (that is a subject does not perform any explicit task). For example, failure to reconfigure one's network in a preparatory phase of an upcoming task leads to more errors (Ekman *et al.*, 2012). And a group of regions whose activity is – on average – strongly correlated while a subject is resting quietly, shows pronounced desynchronization over extended periods of time (Chang & Glover, 2010). Further, spontaneous variability of functional networks has not only been observed in humans but also in anesthesized macaques and rats (Hutchison *et al.*, 2013b; Majeed *et al.*, 2011).

Together these observations suggest that there is no "one" functional network, but, rather, transient connections that form and dissolve. The study of brain network dynamics is only in its early days and there are many unanswered questions. For example, how are these fluctuations best analyzed? What is their functional significance? And what causes their apparent non-stationarity?

Many insights into the average functional organization of the human brain were made possible by methodological progress. New techniques are now needed to characterize dynamic brain networks. In this thesis we aim to adapt tools from signal processing, such as multivariate decompositions and wavelets, to study network dynamics that are driven by changes in a subject's external environment or that occur spontaneously. Why should these tools be used? Multivariate decompositions find characteristic spatiotemporal properties in the data, which makes them excellent tools for exploratory data analyses. This property is especially relevant for the study of spontaneous network variability. Wavelets can provide a view of the data from global to local spatial scales. This seems instructive for fMRI data because brain networks have both hierarchical and modular structure: correlations between brain systems can be further subdivided into correlations between regions within these systems (Ferrarini *et al.*, 2009; Meunier *et al.*, 2009b; Doucet *et al.*, 2011) Thus, we expect co-activations to display different functional characteristics and interactions at different spatial scales. Of particular interest is how classical wavelet analysis can be adapted to the complex and evolving topology of functional brain networks. These tools will be used to gain insight into the functional relevance of network fluctuations and to study how external and internal factors modulate functional interactions.

1.2 Organization and contributions

This dissertation is organized following a paper-based style. In introductory Chapter 2 we present an overview of how brain function and organization can be studied with fMRI. This chapter serves as a background for the published or submitted articles, which are reproduced in Chapters 3 to 5, and for the unpublished manuscript of Chapter 6. In what follows we discuss the main objectives and contributions of each chapter:

Chapter 3: Co-varying functional interactions
How can we identify recurring network configurations over time?
How does multiple sclerosis affect dynamic functional interactions?

If functional connections vary independently from each other over time, then no meaningful temporal covariance exists. However, if some functional connections do not vary independently from each other over time, then these connections co-vary. To identify whether there is meaningful covariance among dynamic functional connections, and to reveal groups of co-varying connections if they exist, we propose to employ principal components analysis (PCA). PCA analyzes the covariance of each dynamic connection with every other connection and finds vectors (principal components) that capture maximal covariation. Each component is a linear combination of the original data such that strong co-variance is emphasized. The principal components span a low-dimensional space in which high-dimensional network dynamics unfold and expose underlying structure.

We adapt the use of PCA to dynamic brain networks and validate it by analyzing dynamic brain networks of both healthy subjects and relapse-remitting multiple sclerosis patients. We estimate both the principal components or "FC patterns", which we interpret as basic building blocks of functional networks, and their strengths of expression over time. While the building blocks are common to all

subjects, their expression is subject-specific and we report altered dynamics of some of the most important (in terms of explained covariation) building blocks in this group of patients.

Chapter 4: Interaction of FC patterns

All at once or one by one. Or: Are FC patterns jointly or exclusively expressed?

Here, we analyze how well dynamic brain networks are characterized by either the joint or exclusive contribution of FC patterns. PCA, as proposed in Chapter 3, models the data as a linear combination of principal components, which corresponds to a joint contribution of all FC patterns. As opposed to this, Allen *et al.* (2014) identified "FC states" by temporally clustering dynamic brain networks. In clustering, a dataset is separated into multiple distinct groups (called clusters), which corresponds to the exclusive contribution of one FC pattern.

We address this question by formulating k-means clustering as a sparse matrix factorization, akin to PCA, but subject to different constraints on the temporal expression of the FC patterns. Then, we analyze simulated and experimental fMRI data. FMRI data was acquired both while subjects performed self-driven tasks, such as silently singing a song, and while they rested quietly. During subject-driven tasks, dynamic networks were well described by the contribution of a single FC pattern. During resting-state, the joint expression of several FC patterns better described the data. These results suggest that combinations of FC patterns provide a more meaningful description of dynamic brain networks during resting-state.

Chapter 5: Multi-scale representation of brain activity

Can local and global co-activations efficiently distinguish cognitive states?
How can we identify recurring network configurations in task data?

In this chapter we study instantaneous co-activations of brain regions at multiple scales: we want a close-up view to reveal activity of individual regions, and increasingly coarser views to reveal coherent activity of functionally connected brain regions. We rely on graph wavelet analysis to obtain such a network-guided multi-scale view. Classical wavelet analysis is the mathematical tool that provides a multi-scale view on Euclidean grids (e.g. time for one-dimensional and image coordinates for two-dimensional data), and recent advances in signal processing have generalized wavelets to more complicated, arbitrary underlying topologies. We extend one recent proposal (Hammond *et al.*, 2011) to dynamic networks by relying on a tensor decomposition, which is the higher-order analogue of a matrix decomposition. In the example of alternating cognitive states, we show its ability to identify task-related changes in network dynamics.

As functional interactions are modulated by changes in the external environment, we hypothesized that a wavelet adapted to a resting-state brain network would efficiently represent global co-activations during rest but not during an experimental task. Initial results suggested that the wavelet representation prominently highlights alternating cognitive states. However, using machine-learning techniques, the two states were equally well predicted using either the wavelet or original representation (i.e., regional BOLD activity). This result and alternative strategies are further discussed in the chapter.

Chapter 6: Co-activating brain regions

How do spontaneous patterns of co-activation change in healthy aging?

This chapter is dedicated to the study of age-related changes in the spontaneous variability of brain networks. In the previous chapters we estimated dynamic networks using sliding-window correlation analysis: Pairwise correlations between regional brain activity were estimated for multiple, possibly

overlapping, temporal segments of the scanning session instead of across the entire scanning session. This estimation is intrinsically variable and limits the temporal sensitivity to network changes to the length of the temporal segment. Here, we approximate the FC patterns with rank-one matrices. This simplifies the problem to estimating patterns of co-activations instead of patterns of correlations. We employ temporal clustering (Liu & Duyn, 2013) to separate fMRI data into several distinct co-activation patterns (CAPs). Each CAP describes a unique spatial pattern of frequently co-activating regions.

Healthy aging has consistently been associated with large-scale disruptions of average brain networks (Ferreira & Busatto, 2013). By comparing healthy young and old adults, we confirm these previous findings. Further, to understand whether changes in dynamic properties may underlie such differences, we identify multiple distinct CAPs in elderly subjects. This reveals that elderly subjects engage a core resting-state network less frequently than young subjects, but engage other sets of regions more frequently. The increased occurrence of one of these CAPs in individual subjects was related to worse performance in an episodic memory task. These results suggest that dynamic properties explain some of the commonly observed differences between young and old adults, and the association with memory indicates the functional relevance of dynamic network properties.

Chapter 7 concludes the thesis, integrates the results and presents suggestions for future research. Because each journal article is followed by its own discussion and an outlook, perspectives for future research focus mainly on Chapter 6.

So that other researchers can reuse and build upon methods proposed in this thesis, Matlab scripts have been made available at `http://miplab.epfl.ch/software`.

2 Probing the functional organization of the human brain with fMRI

WE PERCEIVE AND INTERACT with our surroundings and other human beings, we feel sad or happy, we remember. These are some of the most fascinating functions of the human brain. Understanding how our brain supports these and other functions has proven both technically and conceptually challenging. Much progress in understanding the structural and functional organization of the brain has been made in the last one hundred and fifty years. Phrenology, developed by Franz Joseph Gall in the 18th/19th century, proposed that (i) different areas in the brain have different functions, and (ii) a person's personality traits cause detectable bumps and valleys on a person's scalp. While this second idea has been discredited in the 20th century, the idea of functional specialization has been highly influential. Localization of mental functions in specific brain areas was revolutionized by brain lesion case studies. In 1861, Paul Broca described two patients who had lost their ability to produce meaningful speech and who had pronounced brain damage in their left lateral frontal lobe, an area known today as "Broca's area" (Fig. 2.1). This and other case studies since have been crucial, but the patients' brains could only be studied *post mortem* in autopsies.

The development of non-invasive brain imaging techniques, such as electroencephalography (EEG), and more recently, of positron emission tomography (PET) and fMRI, have made it possible to study a wealth of complex brain functions, not only in patients with neurological diseases but also in healthy human subjects. Electroencephalography (EEG) records electrical activity in the brain by placing multiple electrodes on a subject's scalp. EEG picks up electrical signals that are the result of neural

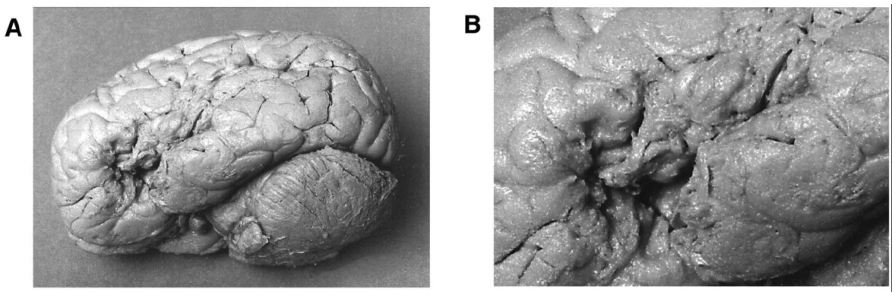


Figure 2.1 – Photographs of the brain of one of Broca's patients. Reprinted from Dronkers *et al.* (2007) with permission from Oxford University Press.

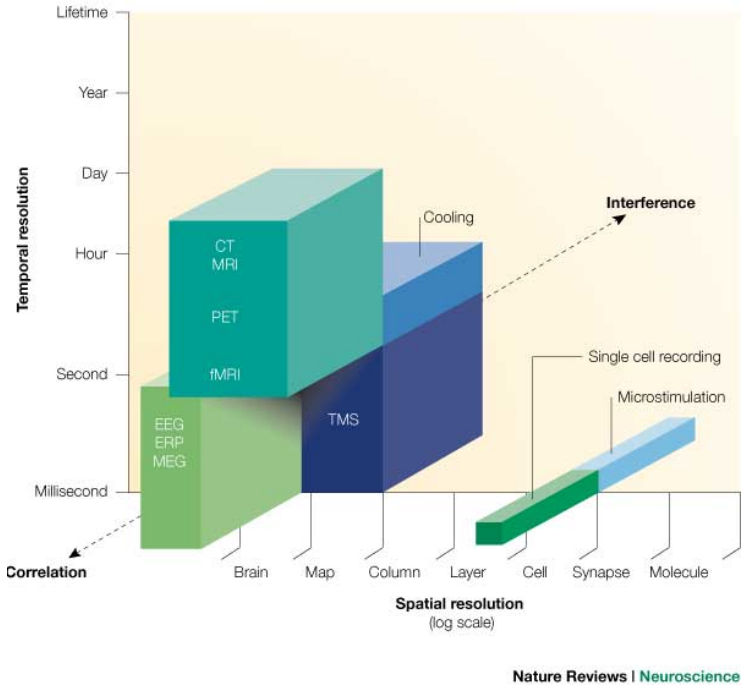
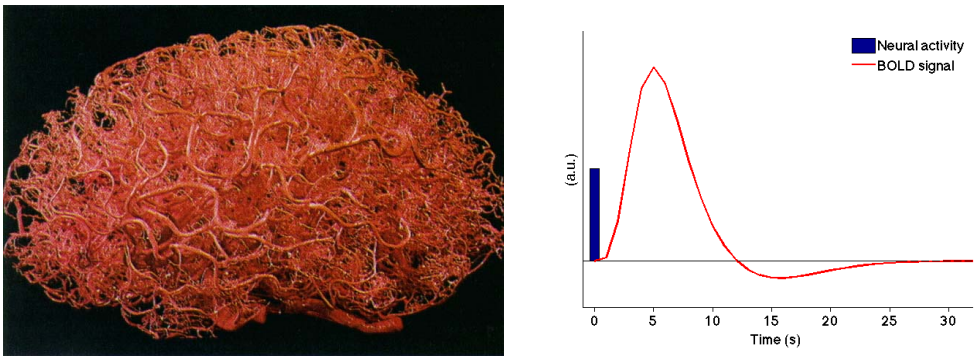


Figure 2.2 – Comparison of temporal and spatial resolutions of imaging and brain stimulation techniques. Techniques with sub-second temporal resolution measure neural activity, and those with supra-second resolution measure delayed correlates of neural activity, such as changes in blood flow. Most techniques provide correlational evidence by testing the association between behavior and brain activity. Trans-cranial magnetic stimulation (TMS) and microstimulation provide causal evidence by transiently interfering with brain activity. CT, computerized tomography; ERP, event-related potential; MEG, magnetoencephalography. Reprinted from Walsh & Cowey (2000) with permission from Nature Publishing Group.

activity throughout the brain. EEG can measure changes in brain activity over split seconds of time, but because it is a surface measurement information about the activity of specific regions is difficult or impossible to obtain, especially when a large number of brain regions are simultaneously active (Fig. 2.2). In PET, changes in cerebral blood flow or cerebral metabolism are detected by scanning a subject after a radioactive-labeled substance, $H_2[\,^{15}O]$ or $[^{18}F]$-fluorodeoxyglucose, respectively, is injected into the blood stream. The amount of radioactive decay measured indicates local cerebral blood flow or glucose metabolism, respectively. The use of PET has, however, declined with the development of fMRI. While PET has the advantage of providing an absolute measure, fMRI has better spatial and temporal resolution and because it uses no radioactivity the same subject can be scanned many times (Fig. 2.2).

2.1 FMRI technique and preprocessing

This section provides a brief introduction to fMRI and we refer to Huettel *et al.* (2004) for a comprehensive overview. FMRI provides an indirect measurement of neural activity because it measures changes in cerebral blood flow and blood oxygenation level. Local neural activity is tightly linked to local changes in blood flow and oxygenation through a phenomenon called "neurovascular coupling": Active or firing neurons consume glucose and oxygen and (through a complex and not yet completely



(a) (b)

Figure 2.3 – (a) Blood vessels in the human brain. Vessels were visualized by filling them with a plastic emulsion and then by dissolving the parenchymal tissue. Reprinted from Zlokovic & Apuzzo (1998) with permission from Wolters Kluwer Health. (b) Model temporal response of the fMRI BOLD signal to neural activity with a sampling time of 1 second.

understood cascade of events) cerebral blood flow increases locally to meet the increased demand. Oxygen is supplied in an intricate and extensive web of vessels, spaced by as little as 20-40 μm in gray matter (GM)[1], and is carried by a protein in red blood cells called hemoglobin (Fig. 2.3a; Zlokovic & Apuzzo, 1998). More oxygen is supplied than consumed, leading to an increase in the ratio of oxy- to deoxyhemoglobin in the blood. Because oxygenated and deoxygenated hemoglobin have different magnetic properties, changes in this ratio cause changes in the MR signal — hence the term "blood oxygenation level-dependent" (BOLD) signal. Since a vascular change, rather than neural activity is measured, the temporal information is delayed and spread out over time. Specifically, changes in the BOLD signal lag neural activity by 1-2 seconds, peak at about 4-6 seconds and then slowly return to baseline levels after another 5-10 seconds (Fig. 2.3b).

In fMRI a temporal series of three-dimensional images is acquired by sampling the BOLD signal every couple of seconds (typically every 1-3 seconds) at tens to hundreds of thousands of locations throughout the brain. These locations are called voxels (volumetric pixels), and typically have an edge length of 1-5 mm and contain millions of neurons. The whole-brain coverage of fMRI is a big advantage and its spatial resolution is excellent (Fig. 2.2). The size of the voxels is thought to approximately correspond to the spatial extent of the vascular response to local neural activity (Pasley & Freeman, 2008). The simultaneous measurement of thousands of voxels and the small magnitude of changes in the BOLD signal make fMRI data very challenging to analyze.

Changes in the BOLD signal cannot only be caused by neural activity, but also by other physiological signals, such as cardiac and respiratory signals, and by non-neuronal sources, such as scanner noise and head motion. Therefore, a number of important steps are taken first to reduce image artifacts and prepare fMRI data for further analysis. These include alignment of functional scans to correct for head motion, and removal of low-frequency drifts by high-pass filtering. To further reduce the impact of non-neuronal contributions, motion parameters from the realignment step and signals from white matter (WM) and cerebrospinal fluid (CSF), which are assumed not to reflect neural activity, can be regressed out from the data. Only the residuals are then retained for further analysis. Alternatively, the nuisance signals can be included as signals of no interest in the analysis (see following sections for

[1] GM consists of neuronal cell bodies, while white matter mostly contains axons or nerve fibers that link GM areas. The names stem from the color of the two tissues.

more details). Low-pass filters, e.g. with cut-off frequency at about 0.1 Hz, are also commonly applied to data acquired in the absence of a task to reduce high-frequency noise and because low-frequency fluctuations are coherent in characteristic patterns (see Section 2.3).

Finally, it is crucial to identify effects that are consistent across subjects. To be able to compare results across subjects it is necessary to bring each subject's data into a common space, a process called "normalization". In this new space, single-subject voxel or region-wise results can be directly compared. However, because of inter-subject variability in the brain's anatomy (and to improve the signal-to-noise ratio) fMRI data is also frequently spatially smoothed.

2.2 Mapping brain functions with task fMRI

Task fMRI aims to localize mental functions in specific brain areas by identifying which voxels are engaged by a behavioral task that represents the function of interest. Thus, the experimenter first needs to design an appropriate behavioral task and associated experimental paradigm. Because the absolute value of the BOLD signal has no meaning, only the *relative* change in the signal, that is an increase or decrease, is of interest. Thus, the experiment must be designed in a way that allows contrasting the BOLD signal during the task with a suitable control condition. Most task fMRI studies use a block design with alternating epochs of task and control conditions, each several scans long. For example, to map motor functions, the subject may be asked to repeatedly move the fingers of one hand and to rest in between. Regions with different BOLD signal during the two conditions are likely to be involved in the brain function under study. If a resting period is used as the control condition, changes in BOLD signal may also be related to more general effects, such as eye movement or attention. More complex designs with cleverly devised control conditions permit to study more high-level or specialized brain functions.

Next, the voxels whose BOLD signal is either significantly increased or decreased by the task, need to be identified. Because of the hemodynamic lag and the influence of nuisance variables, simply subtracting the average BOLD activity of the two conditions is not a sensitive approach. A more powerful approach is multiple linear regression, where the dependent variable $\mathbf{y}$ (the BOLD signal over time of the voxel) is assumed to be related to multiple independent variables $\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_N$ (Friston *et al.*, 1995). The ideal task activation is modeled by $\mathbf{x_1}$, and nuisance variables, such as head motion and signals from the WM and CSF, by $\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_N$. The general linear model (GLM) is given by $\mathbf{y} = \beta_0 + \beta_1\mathbf{x}_1 + \beta_2\mathbf{x}_2 + ... + \beta_N\mathbf{x}_N + \epsilon$, where $\beta_0, \beta_1, ...$ are the parameters, and ϵ models noise. In matrix form $\mathbf{y} = \mathbf{X}\beta + \epsilon$, where $\mathbf{X}$ is the design matrix containing the experimental design, a constant regressor and nuisance variables and β contains all parameters (Fig. 2.4). The parameters are estimated so as to minimize the residual sum of squares $||\mathbf{y} - \mathbf{X}\hat{\beta}||_2^2$, where $\hat{\beta}$ contains the parameter estimates. For invertible $\mathbf{X}^T\mathbf{X}$ and under the assumption of independently and identically normal distributed noise, the solution is given by the ordinary least squares estimator $\hat{\beta} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{y}$.

Whether a voxel is significantly more active during a task is determined with a one-sample t-test of its $\hat{\beta}$ parameter relative to the error of the model. This "activation" test can also be seen as testing the significance of a simple linear combination (i.e., contrast) of the parameters: $0\cdot\hat{\beta}_0 + 1\cdot\hat{\beta}_1 + 0\cdot\hat{\beta}_2 + ... + 0\cdot\hat{\beta}_N$. Testing for "deactivation" corresponds to the linear combination $0\cdot\hat{\beta}_0 - 1\cdot\hat{\beta}_1 + 0\cdot\hat{\beta}_2 + ... + 0\cdot\hat{\beta}_N$. Using this perspective, more complicated questions about the data can be asked by specifying an appropriate linear combination of the $\hat{\beta}$ parameters.

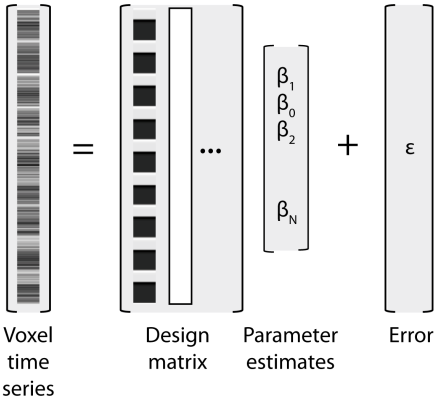
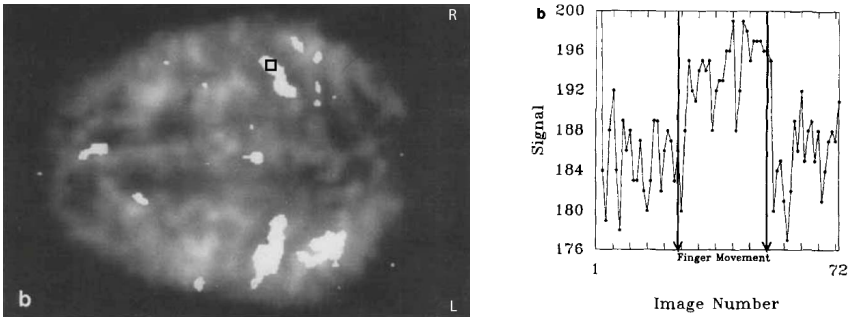


Figure 2.4 – Illustration of the general linear model $\mathbf{y} \approx \mathbf{X}\hat{\beta} + \hat{\epsilon}$.



(a) **(b)**

Figure 2.5 – (a) Brain activity in bilateral primary motor and sensory regions is increased in response to bilateral finger movement, overlaid on an anatomical image. (b) BOLD signal over time from the boxed voxel in (a) (TR= 3 sec). Bilateral finger movement occurred between scans 25 and 51. Adapted from Bandettini *et al.* (1992) with permission from John Wiley and Sons.

The linear regression is estimated for each voxel in turn ("mass-univariate" approach). This leads to a huge number of individual tests, which must be corrected for to avoid an increase in the number of false positive results. Sophisticated methods based on Gaussian random field theory have been developed, which take the spatial correlation of the data into account (Worsley *et al.*, 1996).

One of the first task-based fMRI studies looked at changes in brain activity in response to a simple motor task (Bandettini *et al.*, 1992). The subjects were asked to touch their thumb with each finger in turn, with either one or both hands. Discrete clusters of voxels in the contralateral or bilateral motor cortex were selectively activated by finger movements with one or both hands, respectively (Fig. 2.5). Thousands of task-based fMRI studies have since been published, which have corroborated and further extended findings from lesion and PET studies. However, we note that fMRI cannot establish the necessity or sufficiency of a region for a specific function. In other words, activity in this region (necessity) and only this region (sufficiency) *causes* a certain behavior. If a region is necessary for a specific function, then disrupting neural activity in this region will impair behavior. For example, a patient with a lesion in a specific area who can no longer speak or transcranial magnetic stimulation that disrupts speech can establish necessity. Whether the identified region is sufficient for speech cannot be deduced from such studies. For example, while Broca's discovery was seminal, he was wrong

to conclude that language function is localized in a single brain region. Rather, language processing involves a whole network of regions. Electrical stimulation of a region that elicits a certain behavior can indicate sufficiency. In contrast to modulating brain activity and measuring behavior, fMRI manipulates behavior and measures brain activity. Testing the association between an experimentally manipulated behavior and brain activity provides correlational, rather than causal, evidence but can be an important stepping stone (Fig. 2.2).

Task fMRI assumes functional specialization (or segregation) and, due to the mass-univariate approach, provides a view of regions activated in isolation. Identifying networks of regions that respond in concert necessitates multivariate techniques, such as multivariate pattern analysis (MVPA). MVPA asks how spatial *patterns* of activity change under different conditions and informs about functional integration. In a seminal study, Haxby *et al.* (2001) presented subjects with eight different image categories (faces, cats, houses, chairs, scissors, shoes, bottles, and nonsense pictures), which each evoked a unique pattern of activations in the ventral visual pathway. The patterns, however, also strongly overlapped. Importantly, the patterns allowed for discrimination among all categories of images to be made, even if voxels with maximal activation were excluded from the patterns (i.e., the most significant voxels in a GLM analysis). Other multivariate techniques that have been used to study interactions between different brain regions are discussed in the next section.

2.3 Mapping large-scale networks with resting-state fMRI

In task studies spontaneous fluctuations of the BOLD signal, that is fluctuations occurring during the control resting condition, were viewed as "noise". This view was, however, called into question by the influential observation that spontaneous brain activity in the motor cortex is highly structured: a network of regions, including but not limited to regions that are activated by motor tasks, eliciting strong temporal correlations during rest (Fig. 2.6b; Biswal *et al.*, 1995). Coherent brain activity while a subject is resting is not limited to the motor cortex: Several distributed and reproducible networks, so-called resting-state networks (RSNs), have been revealed (Fig. 2.6c; Beckmann *et al.*, 2005; Damoiseaux *et al.*, 2006).

Several techniques exist for the analysis of coherent spontaneous brain activity. A simple and widely used technique is **seed-correlation analysis**. A reference time course is extracted by averaging the signal in a region of interest (ROI, the seed region) and compared pairwise with the time course of all other voxels in the brain by computing the temporal correlation. Alternatively, the time course of the other voxel may be used as a regressor in a GLM together with nuisance variables. The somatomotor network shown in Fig. 2.6b was for example estimated using seed correlation. Several other RSNs have been identified by placing seeds in different brain regions, including visual (Lowe *et al.*, 1998), auditory (Cordes *et al.*, 2000), and language networks (Cordes *et al.*, 2000; Hampson *et al.*, 2002) and the default mode network (DMN; Greicius *et al.*, 2003, Fig. 2.6c). The DMN, which has attracted most attention in resting-state studies, comprises a set of brain regions that are deactivated during cognitively demanding tasks but are engaged during internally-oriented cognition, such as remembering and thinking about the future (Shulman *et al.*, 1997; Mazoyer *et al.*, 2001; Raichle *et al.*, 2001; Buckner *et al.*, 2008).

An advantage of seed-correlation analysis is its straightforward interpretation, as the obtained voxel-wise FC maps directly show the regions most strongly connected to the seed. Seed-correlation analysis is essentially univariate because it ignores the relationship among other voxels.

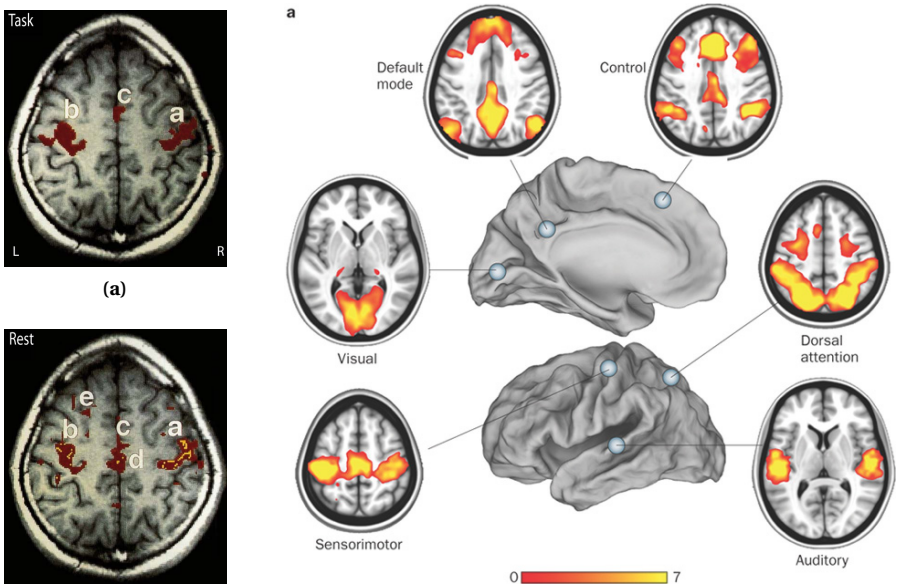


Figure 2.6 – (a) Increased brain activity in a single subject in response to bilateral finger movement, overlaid on an anatomical image. (b) Seed correlation map of a single voxel in the left motor cortex (region b) during rest. Red colored regions indicate $r > 0.35$, yellow ones $r < -0.35$. Reprinted from Biswal *et al.* (1995) with permission from John Wiley and Sons. (c) 6 of the commonly observed seed correlation maps, which were obtained by placing the seed in different brain regions. Reprinted from Zhang & Raichle (2010) with permission from Nature Publishing Group.

Data-driven analyses of spontaneous brain activity, which do not require the definition of a seed region, are another popular choice. For example, principal component analysis (PCA; Friston *et al.*, 1993), and, more commonly, independent component analysis (ICA; McKeown *et al.*, 1998; Kiviniemi *et al.*, 2003; Beckmann *et al.*, 2005; Damoiseaux *et al.*, 2006). ICA factorizes a *voxel* × *time* matrix into spatial maps and associated time courses. Statistical independence of either the spatial maps or time courses is enforced. Spatial independence is the more common assumption for fMRI data, but temporal independence has recently also been explored with promising results (Smith *et al.*, 2012). In spatial ICA, the spatial maps or independent components reveal voxels with coherent activity and are interpreted as RSNs. Fig. 2.7c shows several RSNs that are commonly and consistently identified with spatial ICA and compares them to "activation networks", which were estimated by applying ICA to activation images in a large database of fMRI and PET task studies. Importantly, these networks include the somatomotor network first reported by Biswal *et al.* (1995) (RSN 6), as well as the DMN (RSN 4) and networks related to vision (RSNs 1-3), hearing (RSN 7), and higher-level cognitive processing (RSNs 8-10). An early study using ICA was key in suggesting a role of the DMN in episodic memory: Greicius *et al.* (2004) showed both that the DMN includes the hippocampus and that it is preferentially affected in Alzheimer's, a disease that causes memory loss.

Advantages of ICA are that it is a multivariate method and that several sources of noise are factored out. These components of no interest can for example be identified by their localization in WM or because their time courses are related to head motion. Disadvantages of ICA are that it is an iterative algorithm, which can get stuck in local optima, resulting in variability in repeated runs of the algorithm,

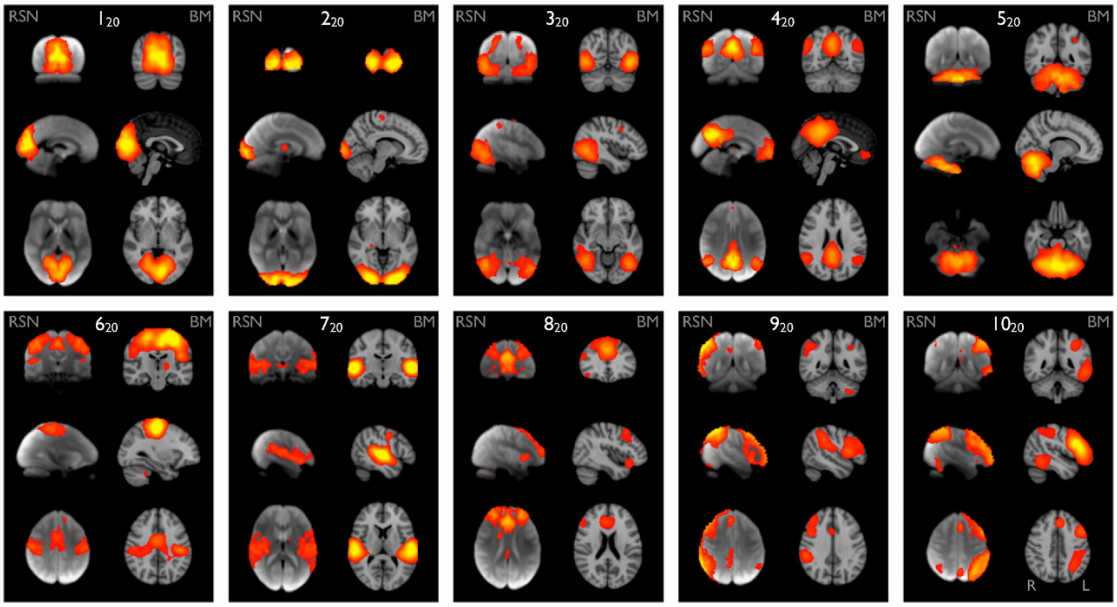


Figure 2.7 – 10 resting-state networks (RSN) and well-matched "activation networks" (in the figure denoted with BM for BrainMap database) estimated from a database with results from task fMRI studies. Reprinted from Smith *et al.* (2009) with permission.

and the necessary model order selection, which remains somewhat arbitrary. However, for multiple, interacting underlying brain systems it is likely that there is no single, best model order (Cole *et al.*, 2010). In ICA for resting-state fMRI it has become common to estimate 20 components, even though higher-dimensional ICAs have also been used to fractionate RSNs into smaller sub-networks (Smith *et al.*, 2009). The higher the number of estimated components, the more difficult it can of course become to identify the relevant ones.

Seed-correlation analysis and ICA both identify large-scale networks of voxels with coherent activity and generally lead to similar results. However, they do not inform about between-network interactions. For ICA, such interactions can be estimated by correlating the time courses of different components (Jafri *et al.*, 2008). When both within- and between-network interactions are of interest, it has been more common to estimate **functional brain networks or graphs** (Fig. 2.8a). Firstly, ROIs, usually referred to as nodes, are defined. They can be defined either functionally, e.g. with high-dimensional ICA or spatial clustering (Craddock *et al.*, 2012; Allen *et al.*, 2014), or anatomically with an atlas, such as the Automated Anatomic Labeling (AAL) atlas (Tzourio-Mazoyer *et al.*, 2002). A small number of studies have also kept the data in voxel space. How the brain is parcellated into regions will affect the estimation of functional interactions and must be chosen by the researcher. Secondly, a representative time series is estimated for each region. For example, by spatially averaging the time courses of all voxels within the region or by estimating the first eigenvariate. Thirdly, the network is estimated from the statistical dependency between all pairs of time courses, such as their temporal correlation or mutual information (Salvador *et al.*, 2005). Brain networks are often visualized in brain space by placing each node at its anatomical coordinate, as shown in Fig. 2.8a. Alternatively, brain networks have been organized circularly or by functional similarity using low-dimensional or force-based embeddings.

The analysis of functional networks with tools from graph theory has for example shown that the brain is a "small-world" network, which promotes both functional segregation and integration (Achard

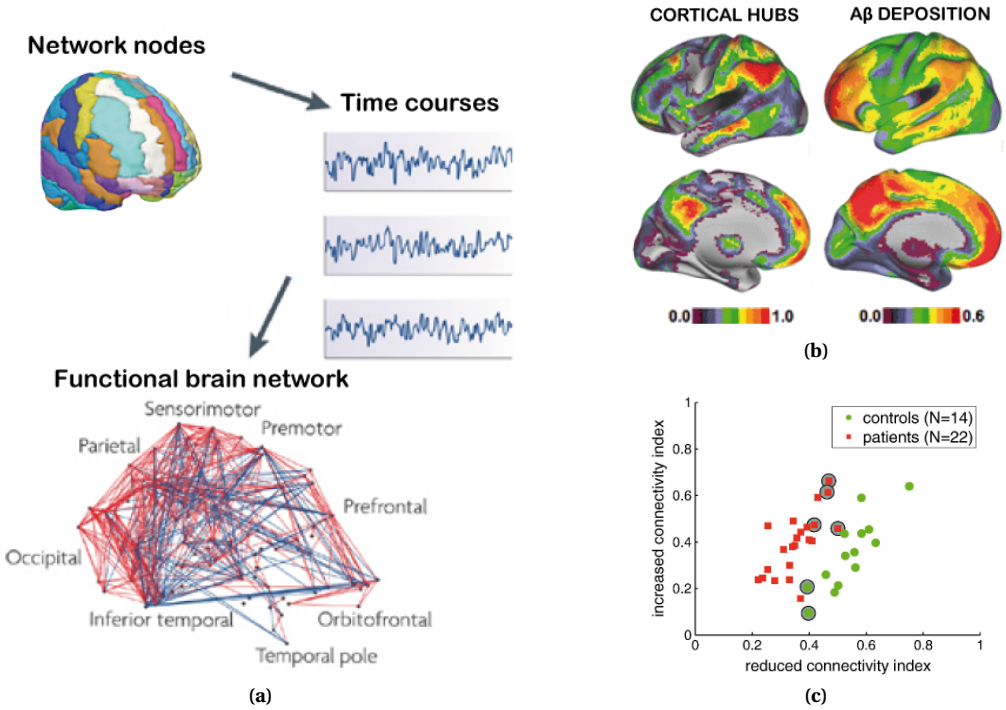


Figure 2.8 – (a) BOLD signal is recorded from multiple anatomically or functionally defined brain regions, and the connections of a functional network are estimated by correlating regional brain activity. Adapted from Bullmore & Sporns (2009) with permission from Nature Publishing Group. (b) Strongly connected voxels in young adults (so-called hubs) show increased amyloid-β deposition in patients with Alzheimer's disease. Adapted from Buckner *et al.* (2009) with permission from Society for Neuroscience. (c) Index of discriminative connectivity alterations, defined as the sum of the feature weights learnt by classifier, after weighing them by the strength of the connection in the subject. Reprinted from Richiardi *et al.* (2012) with permission from Elsevier.

et al., 2006; Bullmore & Sporns, 2009) and that highly connected brain regions, which act as gateways in information processing, are particularly vulnerable to amyloid-β deposition in Alzheimer's disease (Fig. 2.8b; Buckner *et al.*, 2009). Several recent studies have also combined FC analysis and MVPA to analyze how interactions between brain regions are altered by changes in cognitive states or in disease (Craddock *et al.*, 2009; Dosenbach *et al.*, 2010; Richiardi *et al.*, 2011, 2012, 2013; Shirer *et al.*, 2012). Such methods can identify a subnetwork comprised of multiple connections with discriminative ability in single subjects (Fig. 2.8c).

The analysis of spontaneous brain activity with a variety of techniques can thus provide novel insights into the functional organization of the human brain. That spontaneous brain activity is functionally relevant is not only indicated by its non-random nature, but also by several other lines of evidence: (i) The brain at rest consumes 20% of the body's energy but represents only 2% of body mass, and its demand for energy increases by less than 5% during task performance (Fox & Raichle, 2007). (ii) RSNs resemble task-induced activation patterns (Fig. 2.7; Smith *et al.*, 2009). (iii) Brain activity is coherent in various stages of consciousness (Boly *et al.*, 2008; Greicius *et al.*, 2008). (iv) The strength and extent of functional interactions are modulated by preceding cognitive states, such as the performance of a language or learning task and exposure to emotional stimuli (Waites *et al.*, 2005; Albert *et al.*, 2009; Lewis *et al.*, 2009; Eryilmaz *et al.*, 2011). (v) In almost all psychiatric and neurological

disorders, such as in Alzheimer's as mentioned above, patterns of intrinsic activity are disrupted (for reviews see e.g. Greicius, 2008; Fox & Greicius, 2010; Menon, 2011). (vi) Spontaneous activity accounts for variability in task BOLD responses and human perception and behavior: For example, the extent of lateralization of a seed correlation map of a region in the left somatomotor cortex predicts the lateralization of the subject's finger tapping response (De Luca *et al.*, 2005) and spontaneous activity contributes to inter-trial variability of brain activity in response to the same stimulus (Fox *et al.*, 2006). Prestimulus spontaneous activity of regions within several RSNs affects somatosensory, auditory and visual perception (Boly *et al.*, 2007; Hesselmann *et al.*, 2008; Sadaghiani *et al.*, 2009), as well as vigilance (Thompson *et al.*, 2013a). And networks reflecting emotional salience processing and cognitive control correlate with prescan anxiety and executive task performance, respectively (Seeley *et al.*, 2007).

What the purpose of spontaneous activity is, is still debated, however. Based on theoretical, computational and empirical studies, several potential (non-exclusive) functions have been proposed: (i) spontaneous activity could simply be "noise" shaped by the underlying non-random connectivity, (ii) ongoing activity could enable quicker responses to future events, and in a similar vein, reflect predictions about future use, (iii) it may serve to selectively process previous experiences and consolidate learned information, and (iv) it may allow the exploration of possible functional interactions (Fox & Raichle, 2007; Sadaghiani *et al.*, 2010; Deco *et al.*, 2013; Harmelech & Malach, 2013).

3 Temporally co-varying functional connections during rest in healthy adults and multiple sclerosis patients

WHERE AND WHEN THE ACTIVITY OF BRAIN REGIONS is synchronous informs about a subject's biological age (Dosenbach *et al.*, 2010), how well a novel motor skill is mastered (Bassett *et al.*, 2011), whether a visual discrimination task is successfully performed (Ekman *et al.*, 2012), or a subject's cognitive state (Richiardi *et al.*, 2011; Shirer *et al.*, 2012). Contrary to these studies, which revealed the pronounced flexibility of brain networks in response to changing environmental conditions, networks estimated during resting-state acquisitions have typically been considered to be constant throughout the duration of a scan. While average or "static" FC analyses have contributed significantly to our understanding of the human brain in terms of large-scale organization into networks, recent work has highlighted significant, and potentially clinically relevant, deviations from static FC over time (Fig. 3.1, e.g. Chang & Glover, 2010; Hutchison *et al.*, 2013b). These spontaneous fluctuations in FC were previously ignored, but "dynamic FC" (dFC) has become a major new research direction in neuroimaging. Several authors have proposed novel approaches to study the spontaneous evolution of brain networks, which include generalizations of seed-correlation (Kang *et al.*, 2011) and ICA, based on using short temporal segments of the data (Sämann *et al.*, 2011; Kiviniemi *et al.*, 2011), detection of "typical" brain activity through serial spatial correlations with a template or through temporal clustering (Majeed *et al.*, 2011; Liu & Duyn, 2013), and the analysis of windowed temporal correlations by visual inspection and temporal clustering (Hutchison *et al.*, 2013b; Li *et al.*, 2014; Allen *et al.*, 2014). The first four studies have revealed whole-brain or seed-based patterns of *co-activated* voxels, and the last three studies are based on windowed correlations and have revealed reproducible *network configurations* across time.

In this work we focus on dFC estimated from windowed correlations, but rather than separating dFC into states as in Li *et al.* (2014) and Allen *et al.* (2014), our goal is to identify underlying interaction patterns that act as basic building blocks for the evolution of dFC. To identify these patterns we make use of principal component analysis (PCA), a workhorse of exploratory multivariate data analysis, dimensionality reduction and data preprocessing (Jolliffe, 1986). PCA creates new, orthogonal (and hence decorrelated) components, which are linear combinations of the original variables and most

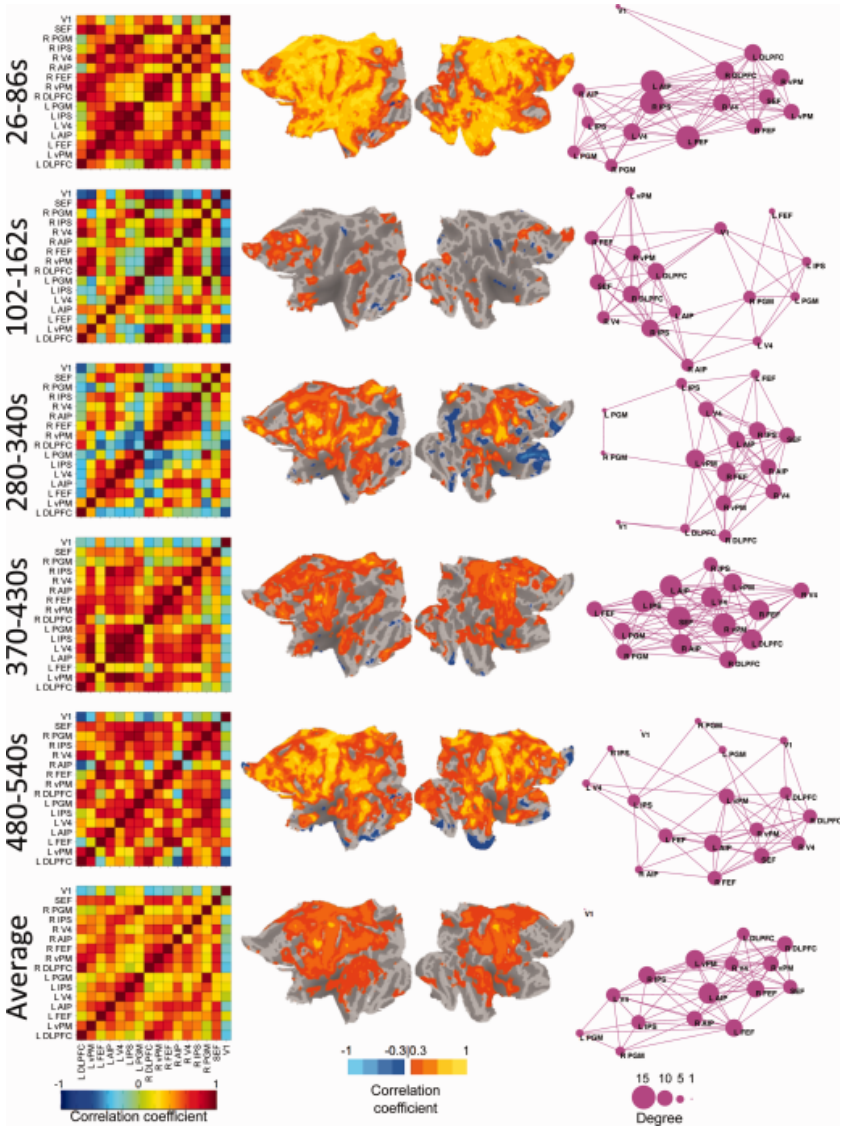


Figure 3.1 – Functional networks change across short time scales. Several networks are estimated in a human subject from 1 minute of data (rows 1 and 5) and differ from a network estimated across the entire duration of the scan. The left column shows pairwise correlation matrices between 16 brain regions, center columns show left frontal eye field (L FEF) correlation maps on a flattened cortical surface, and the right column shows a graph representation of the network. Reprinted from Hutchison *et al.* (2013b) with permission from John Wiley and Sons.

efficiently represent the variance in the data. Because high-dimensional data contains redundant information, a majority of the variance is generally concentrated in the first few components, which then define a new, low-dimensional space. Identifying which connections have a large contribution to each component can reveal underlying structure as these connections likely co-vary over time. Fig. 3.2 gives a conceptual overview of our proposed analysis pipeline.

The second goal of this chapter is to track the temporal expression of the connectivity patterns in healthy subjects and minimally-disabled relapse-remitting multiple sclerosis (RRMS) patients. Notably, MS and many other neurodegenerative diseases are known to disrupt the (average) large-scale, resting-

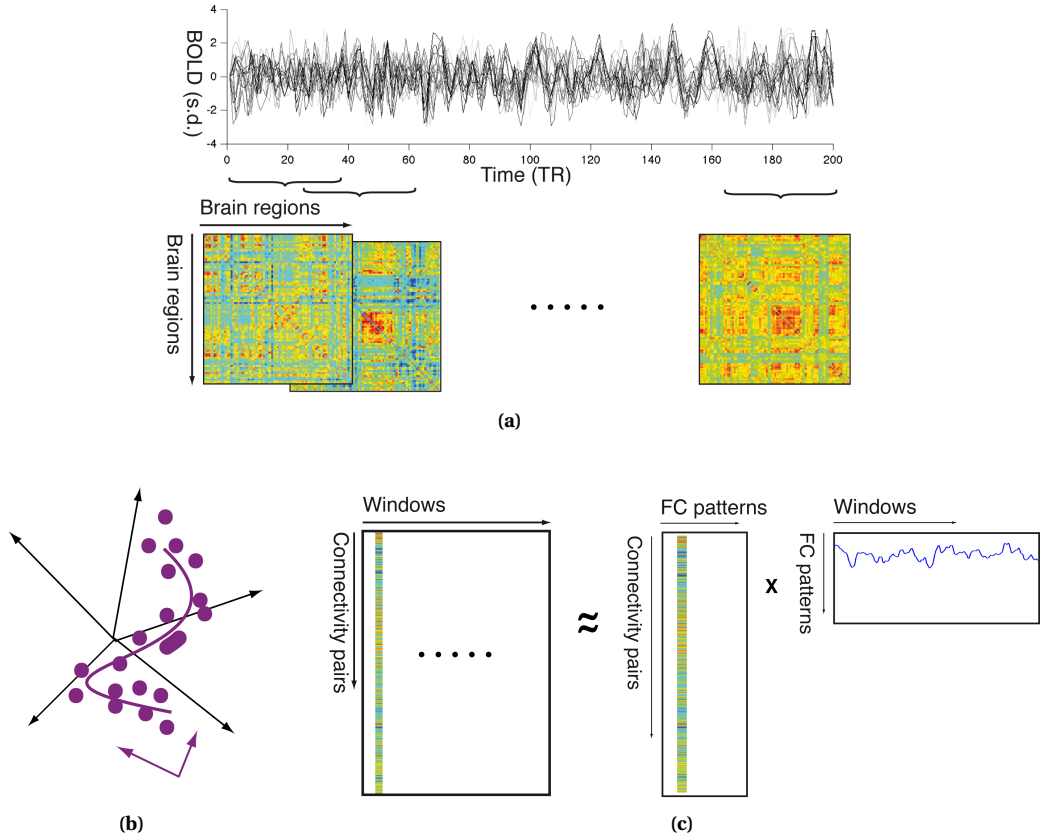


Figure 3.2 – (a) dFC is estimated from sliding-window correlations of regional BOLD activity. (b) High-dimensional dFC estimates (circles) follow a trajectory, indicated by the line, that lies in a lower-dimensional space. (c) Building blocks of dFC are identified using PCA: The axes of the lower-dimensional space are the FC patterns and the coordinates of dFC in this space track their temporal expression.

state networks of the human brain (e.g. Lowe *et al.*, 2002; Greicius *et al.*, 2004; Seeley *et al.*, 2009; Fox & Greicius, 2010; Richiardi *et al.*, 2012). While MS is a heterogeneous disease, previous work by our lab showed that individual subjects could be reliably classified as either healthy subjects or RRMS patients based on their static FC (sensitivity 82%, specificity 86%; Richiardi *et al.*, 2012). Here, we re-analyze the same data set using a different perspective. Static FC, however, provides an average view of resting-state networks and hides the fact that connectivity patterns of any single brain region may vary over time. Important dynamic properties, potentially including dynamic compensatory adjustments, may thus remain hidden (for a recent review see also Hutchison *et al.*, 2013b). FC dynamics have the potential to reveal novel insights into how resting-state networks in the human brain are targeted by neurodegenerative diseases. In this chapter we present early evidence that dynamic properties are altered in multiple sclerosis.

3.1 Journal article

Published in: NeuroImage, 2013; 83:937-950.

Principal components of functional connectivity: A new approach to study dynamic brain connectivity during rest

Nora Leonardi,[1,2] Jonas Richiardi,[3,4] Markus Gschwind,[4,5,6] Samanta Simioni,[5] Jean-Marie Annoni,[4,6,7] Myriam Schluep,[5] Patrik Vuilleumier,[4,6] and Dimitri Van De Ville[1,2]

[1] Medical Image Processing Lab, Institute of Bioengineering, École Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland; [2] Medical Image Processing Lab, Department of Radiology and Medical Informatics, University of Geneva, Geneva, Switzerland; [3] Department of Neurology and Neurological Sciences, Stanford University, Stanford, USA; [4] Department of Neuroscience, University Medical Center (CMU), University of Geneva, Geneva, Switzerland; [5] Department of Clinical Neurosciences, Centre Hospitalier Universitaire Vaudois (CHUV) and University of Lausanne, Lausanne, Switzerland; [6] Department of Neurology, Hopitaux Universitaires de Genève (HUG), Geneva, Switzerland; [7] Neurology unit, University of Fribourg, Fribourg, Switzerland

Functional connectivity (FC) as measured by correlation between fMRI BOLD time courses of distinct brain regions has revealed meaningful organization of spontaneous fluctuations in the resting brain. However, an increasing amount of evidence points to non-stationarity of FC; i.e., FC dynamically changes over time reflecting additional and rich information about brain organization, but representing new challenges for analysis and interpretation. Here, we propose a data-driven approach based on principal component analysis (PCA) to reveal hidden patterns of coherent FC dynamics across multiple subjects. We demonstrate the feasibility and relevance of this new approach by examining the differences in dynamic FC (dFC) between 13 healthy control subjects and 15 minimally disabled relapse-remitting multiple sclerosis patients. We estimated whole-brain dFC of regionally-averaged BOLD activity using sliding time windows. We then used PCA to identify FC patterns, termed "eigenconnectivities", that reflect meaningful patterns in FC fluctuations. We then assessed the contributions of these patterns to the dFC at any given time point and identified a network of connections centered on the default-mode network with altered contribution in patients. Our results complement traditional stationary analyses, and reveal novel insights into brain connectivity dynamics and their modulation in a neurodegenerative disease.

3.1.1 Introduction

Spontaneous fluctuations of the functional MRI (fMRI) blood-oxygen-level-dependent (BOLD) signal are not random but temporally coherent between distinct brain regions. While these fluctuations were long considered as "noise", Biswal et al. (1995) showed that fluctuations of motor areas were correlated even in the absence of a motor task. Several other networks of coherent BOLD activity between remote brain regions have since been identified, including visual, auditory, language and attention networks, and a network called the "default mode network" (DMN) which reduces its activity during attention-demanding tasks. These networks of regions with coherent activity during rest are

consistent across subjects and closely resemble the brain's functional organization of evoked responses (Damoiseaux *et al.*, 2006; Fox & Raichle, 2007; Smith *et al.*, 2009; Laird *et al.*, 2011). Coherent BOLD activity persists during sleep and in anaesthetized monkeys, suggesting that it reflects a fundamental property of the brain's functional organization (Larson-Prior *et al.*, 2009; Vincent *et al.*, 2007).

Coherent BOLD activity, known as "functional connectivity" (FC), is modulated by learning (Bassett *et al.*, 2011), cognitive and affective states (Richiardi *et al.*, 2011; Eryilmaz *et al.*, 2011; Shirer *et al.*, 2012; Ekman *et al.*, 2012; Cribben *et al.*, 2012) and also spontaneously (Chang & Glover, 2010; Britz *et al.*, 2010; Kitzbichler *et al.*, 2009). Chang et al. (2010) showed that FC between the posterior cingulate cortex, a key region of the default mode network, and various other brain regions was highly dynamic over time. Standard stationary FC analyses assume temporal stationarity and are blind to the temporal evolution of FC. Several recent studies have since confirmed the non-stationarity of FC (Majeed *et al.*, 2011; Kang *et al.*, 2011; Kiviniemi *et al.*, 2011; Handwerker *et al.*, 2012; Hutchison *et al.*, 2013b; Li *et al.*, 2014). We will use the terms dFC to specifically refer to fluctuating connectivity during rest, and stationary FC to refer to connectivity estimated under the assumption of temporal stationarity. Chang et al. (2013b) reported a link between heart rate variability and dFC in some brain regions related to arousal and vigilance, suggesting a neural basis for these fluctuations, and Hutchison et al. (2013b) showed that these fluctuations persist in anesthesized monkeys in the absence of conscious, cognitive processing.

When no temporal stationarity is assumed, whole-brain and multi-subject FC studies get vastly more complex. Also, given that there is no task, dFC patterns must be learned from the data. Due to these challenges many studies assessing FC changes have focused on individual brain regions or networks, such as the extensively studied DMN. Allen et al. (2014) recently proposed a data-driven approach based on k-means clustering to identify whole-brain "FC states" in the resting brain. The authors showed that these states differed strongly from FC patterns identified in stationary analyses and suggested that they represent departures from average FC patterns as FC between individual brain regions appears or disappears. They also highlighted the benefit of whole-brain studies to understand the dynamic interplay between different brain regions or networks.

Given the increasing evidence of dFC during rest and its importance for characterizing the brain's intrinsic functional organization, the aim of our study was to develop a data-driven technique to reveal characteristic fluctuations of whole-brain FC. Our method exploits whole-brain FC of multiple subjects to identify connectivity patterns with coherent dynamics and their contributions across time. Specifically, we estimated dFC by calculating the correlations between the BOLD activity of anatomically defined brain regions covering the whole brain using sliding time windows (Chang & Glover, 2010). We then temporally concatenated the dFC between all pairs of brain regions across subjects to construct a *connectivity* × *time-subject* data matrix. We then applied principal component analysis (PCA), which finds intrinsic structure in the data that here represents patterns of coherent FC dynamics. Using these FC patterns, termed "eigenconnectivities", the FC at any specific time can be represented as a weighted sum. Thus, we can interpret the eigenconnectivities as building blocks of dFC with a contribution that varies across time. In contrast to the clustering approach employed in electroencephalography (EEG) microstate (Pascual-Marqui *et al.*, 1995) and FC state analysis (Allen *et al.*, 2014), where stable EEG scalp topographies or whole-brain FC states are identified, we used PCA to identify the most abundant building blocks that reappear across time and subjects.

PCA has previously been applied by Friston et al. (1993; 1994) to voxel-wise PET and fMRI data to determine patterns of correlated brain activity, so-called "eigenimages". In contrast to our work, the authors applied PCA to *activity* and not *connectivity* time courses, and to single subjects only,

which complicates comparisons across subjects. More commonly, voxels with similar resting state activity are identified using spatial independent component analysis (ICA), which decomposes a *voxel* × *time-subject* data matrix into (maximally independent) spatial maps and associated time courses (McKeown *et al.*, 1998; Kiviniemi *et al.*, 2003). The application of PCA and ICA in these studies, however, implicitly assumes FC to be temporally stationary and identifies patterns of voxels with similar resting state activity across the entire scan.

Singular value decomposition[1] (SVD) has also been applied to uncover patterns of relations between brain activity and experimental conditions, behavior or the activity of other voxels in a seed region, in a method called Partial Least Squares (PLS) (McIntosh *et al.*, 1996; Krishnan *et al.*, 2011). The right singular vectors contain patterns of voxels that are associated with e.g. reaction time or seed voxel's activity (i.e., functionally connected). Here, rather than finding relations between the brain activity of different voxels across different conditions, we identify relations between brain *connectivity* time courses in an unsupervised manner.

Stationary FC analysis during rest has been widely applied to understand the functional impact of neurological and psychiatric diseases, such as Alzheimer's disease or schizophrenia, and to search for reliable biomarkers (Fox & Greicius, 2010). To demonstrate the feasibility and relevance of the proposed building blocks, we are interested in early dFC changes in patients with multiple sclerosis (MS). MS is an inflammatory disease that attacks the myelin sheeths in the central nervous system, causing brain lesions, which are randomly distributed in periventricular white matter, but concentrated around the frontal and occipital horns of the lateral ventricle, and at the cortico-subcortical junction (Compston & Coles, 2008). This was also the case in our minimally disabled relapsing-remitting patients (RRMS) group. MS is usually accompanied by sensory or motor deficits and cognitive impairment, but the correlation between clinical disability and focal lesions visible in structural MRI is poor (Barkhof, 2002; Filippi & Agosta, 2010). This is because "hidden" damage occurs in normal appearing brain tissue (NABT), extending well beyond focal lesions (Fu *et al.*, 1998; Hackmack *et al.*, 2012). Alternative imaging methods, such as resting-state fMRI, have started to address the issue of hidden damage and shown disrupted stationary FC. While some studies restricted their analysis to the DMN or motor regions (Lowe *et al.*, 2002; Rocca *et al.*, 2010; Bonavita *et al.*, 2011; Dogonowski *et al.*, 2013), others have reported changes in several large-scale resting state networks, including the sensorimotor and executive control networks (Roosendaal *et al.*, 2010; Hawellek *et al.*, 2011; Richiardi *et al.*, 2012; Rocca *et al.*, 2012). Together, these results are very promising for characterizing the functional pathology of MS, but the relationship between structural damage and functional changes is still poorly understood (Schoonheim & Filippi, 2012).

Here, we study a group of minimally disabled RRMS patients and demonstrate for the first time how dFC is altered in these patients. Specifically, we examine whether the contribution of the identified FC patterns is disturbed in the early stage of MS. We believe that the study of dFC can further contribute to elucidating functional adaptation and reorganization in this pathology.

[1] There is a well-known relationship between SVD and PCA: for centered data $\mathbf{X}$, the left singular vectors of SVD (which is applied to $\mathbf{X}$) equal the principal components of PCA (which is applied to the covariance matrix $\mathbf{X}\mathbf{X}^T$): $\mathbf{X} = \mathbf{U}\Sigma\mathbf{V}^T \rightarrow \mathbf{X}\mathbf{X}^T = \mathbf{U}\Sigma\mathbf{V}^T(\mathbf{U}\Sigma\mathbf{V}^T)^T = \mathbf{U}\Sigma\mathbf{V}^T\mathbf{V}\Sigma\mathbf{U} = \mathbf{U}\Sigma^2\mathbf{U}^T$ since $\mathbf{V}^T\mathbf{V} = \mathbf{I}$. The columns of $\mathbf{U}$ are called left singular vectors or principal components.

3.1.2 Methods

Subjects

Subjects and MRI data were the same as in a previous study of our group (Richiardi *et al.*, 2012). RRMS patients were diagnosed according to McDonald's criteria (Polman *et al.*, 2005) and selected using criteria for 1) mild to moderate neurological disability with no impaired ambulation (Expanded Disability Status Scale (EDSS) ≤ 2.5 (Kurtzke, 1983)), 2) no clinical relapse or corticosteroid therapy during the 6 weeks before inclusion in the study, and 3) no other neurological or psychiatric illness according to DSM-IV criteria. 22 RRMS patients (mean ± SD age 36.8 ± 8 years, 14 females, EDSS 1.9 ± 0.4, disease duration 4.7 ± 3.5 years) and 14 healthy control (HC) subjects (age 38.4 ± 6 years, 9 females) were included. The study was approved by the local institutional ethics committee and all subjects provided informed consent.

Data acquisition

Data were acquired on a Siemens 3T TrioTIM using a 32-channel head coil. Functional imaging data were acquired in one session using gradient-echo planar imaging (TE = 27 ms, TR = 1.1 s, flip angle = 90°, matrix = 64 × 64, 21 transverse slices, voxel size = 3.75 × 3.75 × 5.63 mm^3, 450 volumes). Subjects were instructed to lie still with their eyes closed and not to fall asleep (which was confirmed in a debriefing after the scan).

T1 anatomical scans were collected using an MPRAGE sequence (TE = 2.98 ms, TR = 2.4 s, flip angle = 9°, 160 slices, voxel size = 1 × 1 × 1.2 mm^3).

Data preprocessing

Anatomical and functional data were preprocessed using SPM8 (`www.fil.ion.ucl.ac.uk/spm`), and a combination of in-house MATLAB scripts and scripts from the DPARSFA toolbox (Chao-Gan & Yu-Feng, 2010). The first ten functional volumes were discarded to allow for T1 equilibration effects. The other 440 functional volumes were spatially realigned to the mean image, detrended (linear and quadratic trends) and bandpass filtered ($0.01 - 0.15$ Hz). To further minimise spurious changes in connectivity related to motion, we "scrubbed" our data by removing all volumes that either 1) had a frame-wise displacement $FD = |\Delta x| + |\Delta y| + |\Delta z| + |\Delta\alpha| + |\Delta\beta| + |\Delta\gamma| > 0.5$, where the rotational displacements α, β, γ were converted to millimetres by calculating the displacement on the surface of a sphere of radius 50 mm (Power *et al.*, 2012), or 2) were considered as intensity spikes, defined as a whole-brain intensity of more than 2 standard deviations from the average intensity across the duration of the scan. We also excluded the preceding 2 and following 4 volumes for both indices, corresponding to a similar temporal coverage as in Power et al. (2012). Because of the sensitivity of functional connectivity to motion and to retain sufficient data for the dynamic connectivity analysis, we excluded 1 HC subjects and 7 RRMS patients with either maximal head motion above 3 mm or 2°, or more than 10 out of 440 volumes marked for scrubbing (before adding the 6 surrounding volumes). All results thus reflect a total of 13 HC subjects and 15 RRMS patients.

The functional data were coregistered with each subject's anatomical data. The anatomical data were then segmented (*new segmentation* algorithm of SPM8, an extension of the *unified segmentation* algorithm (Ashburner & Friston, 2005)) and regionally parcellated using the automated anatomical labeling atlas, which divided the brain into 90 anatomical regions of interest (Supplementary Table A.1)

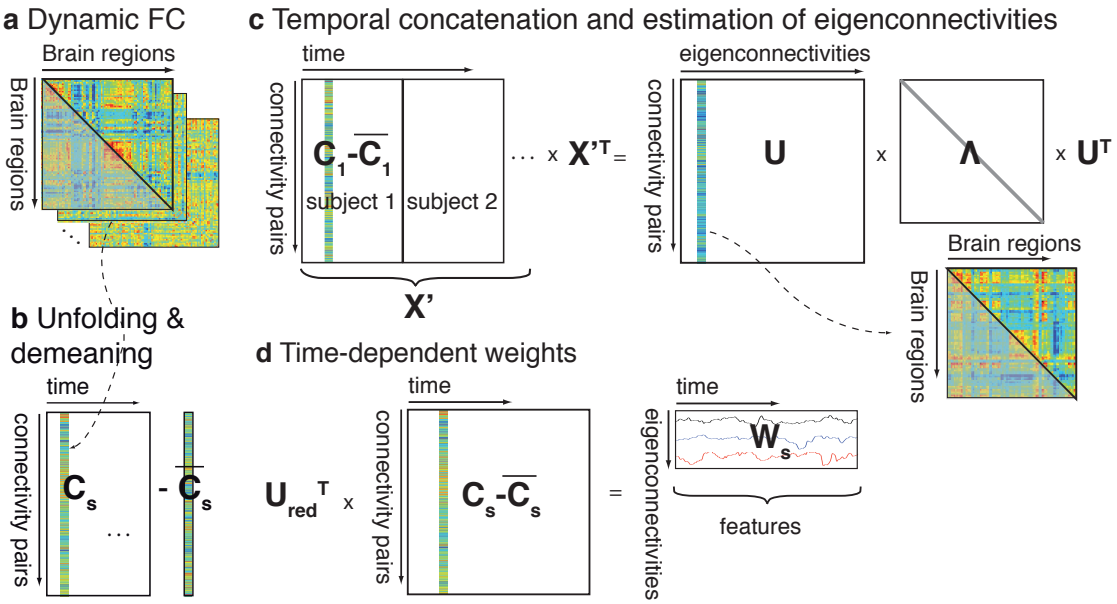


Figure 3.3 – Diagram of the analysis pipeline: (a) dFC between 88 brain regions was computed as sliding window correlations between the activity of all regions for each subject. (b) The upper triangular part of each 88 × 88 correlation matrix was unfolded and concatenated across time to form a dFC matrix $\mathbf{C_s}$. The mean correlation value across time was subtracted from all connectivity pairs (i.e., row-wise centering). (c) The dFC matrices were concatenated across subjects to form $\mathbf{X}'$ to which we applied PCA, yielding eigenconnectivities (columns of $\mathbf{U}$). Eigenconnectivities can be visualized after reshaping them into a 88 × 88 matrix and symmetrizing. (d) A set of time-dependent weights $\mathbf{W}_s$ for each subject was calculated by projecting the centered dFC matrix onto a few eigenconnectivities.

(Tzourio-Mazoyer *et al.*, 2002). The segmentation step provided a deformation field, which was used to warp the structural atlas to the subject's native space. We estimated regional mean time series by averaging the fMRI signal over all voxels in each brain region. The time series from the bilateral globus pallidus were discarded due to ventral signal dropout in some subjects, leaving 88 brain regions. We regressed out nuisance variables from the regional time series (6 head motion parameters, average cerebrospinal fluid from ventricular masks and white matter signal from white matter masks). The global signal was not regressed out as this step is currently debated in the literature.

dFC estimation

We estimated dFC by calculating the pairwise Pearson correlation between the residuals of all 88 brain regions using a sliding-window technique yielding a 88 × 88 correlation matrix for each window (Fig. 3.3a): the sliding-window correlation between the time series x and y was given by $\rho_{xy}(t) = \mathrm{corr}(x[t, t+\Delta t], y[t, t+\Delta t])$, where Δt was the window length in TRs and the window was shifted by two TRs for each estimation. Correlations ρ were Fisher r-to-z transformed ($z = atanh(\rho)$) to make them approximately normally distributed.

We constructed a dFC matrix $\mathbf{C}_s$ or each subject, where each column contained a vectorized correlation matrix (Fig. 3.3b). When vectorizing the correlation matrices, we kept only the upper triangular part due to symmetry, resulting in a $(N^2 - N)/2 \times T_s$ matrix, where $N = 88$ was the number of

brain regions, and T_s the number of windows for subject s. We normalized each subject's dFC matrix by removing the global mean and dividing by its standard deviation (mean and standard deviation over all time points and FC pairs).

The sliding-window technique has previously been used by several other groups to investigate dFC with window lengths varying from 15 to 120 TRs (Chang & Glover, 2010; Bassett *et al.*, 2011; Hutchison *et al.*, 2013b; Handwerker *et al.*, 2012; Allen *et al.*, 2014) and similar step sizes (Handwerker *et al.*, 2012; Hutchison *et al.*, 2013b; Allen *et al.*, 2014). We used window lengths of 30, 40, 60 and 120 TRs, covering 33, 44, 66 and 132 seconds respectively, to investigate the effect of the window length.

Eigenconnectivity estimation

To estimate characteristic connectivity patterns across time and subjects, we applied PCA to a matrix containing the dFC matrices of all HC subjects and RRMS patients; i.e., $\mathbf{X} = [\mathbf{C}_1, \mathbf{C}_2, ...\mathbf{C}_S]$. PCA is a statistical technique to detect unknown, yet structured patterns that explain the differences in the collection of vectorized correlation matrices. A *few* patterns typically explain most of the variation present in *all* of the original correlation matrices, allowing to automatically summarize the data.

Before applying PCA, the mean of each row is normally subtracted from the data: $\mathbf{X}' = \mathbf{X} - \overline{\mathbf{X}}$, where $\overline{\mathbf{X}}$ is a matrix that contains the row-wise means. However, we observed that PCA was partly driven by inter-individual differences in mean correlation levels rather than correlation fluctuations when we subtracted the mean of the concatenated matrix (see section 3.6). Therefore, we subtracted the row-wise means of $\mathbf{C}_s$ for each subject s individually before we concatenated all subjects along the temporal dimension: $\mathbf{X}' = [\mathbf{C}_1 - \overline{\mathbf{C}}_1, \mathbf{C}_2 - \overline{\mathbf{C}}_2, ..., \mathbf{C}_S - \overline{\mathbf{C}}_S]$ (Figs. 3.3b, c).

We then applied PCA, which finds the eigenvectors and eigenvalues of the covariance matrix of $\mathbf{X}'$ (Fig. 3.3c): $\mathbf{X}'\mathbf{X}'^T = \mathbf{U}\Lambda\mathbf{U}^T$, where $\mathbf{U}$ contains orthonormal eigenvectors on its columns and Λ the corresponding eigenvalues on its diagonal; i.e., $\Lambda_{kk} = \lambda_k$. These eigenvectors (or equally principal components) can be seen as features (or spatial modes) that characterize the variation across the collection of correlation matrices. Because they capture connectivity patterns, we call them "eigenconnectivities".

The matrix $\mathbf{U}$ is of size $(N^2 - N)/2 \times (N^2 - N)/2$; i.e., we obtain $(N^2 - N)/2$ eigenconnectivities. However, the eigenconnectivities with the largest associated eigenvalues explain most of the variation in the collection of dFC matrices and a small number can be used to efficiently approximate the data. Approximating the data using only the K eigenconnectivities with the K largest eigenvalues, results in a retained variance of $\sum_{k=1}^{K} \lambda_k / \sum \lambda_k$. The retained variance is equal to the approximation accuracy $1 - ||\mathbf{X}' - \mathbf{X}'_{\text{approx}}||^2 / ||\mathbf{X}'||^2$, where $\mathbf{X}'_{\text{approx}} = \mathbf{U}_{\text{red}}\mathbf{U}_{\text{red}}^T\mathbf{X}'$, and $\mathbf{U}_{\text{red}}$ contains only the first K eigenconnectivities.

Time-dependent contribution of eigenconnectivities

A subject's dFC matrix can be efficiently represented by projecting it onto a few eigenconnectivities: $\mathbf{W}_s = \mathbf{U}_{\text{red}}^T(\mathbf{C}_s - \overline{\mathbf{C}}_s)$, where $\mathbf{W_s}$ is a matrix of size $K \times T_s$ containing the weights of K eigenconnectivities at all T_s windows (Fig. 3.3g). These weights are similar to the time-dependent profiles obtained for the "eigenimages" in Friston et al. (1993), and conceptually related to the subject-dependent time courses calculated in independent component analysis (ICA) by dual regression (Beckmann *et al.*, 2009). In ICA, spatial maps from group-ICA are used as regressors to estimate a time course for each independent component and subject. Here, we use the eigenconnectivities obtained from group analysis as regressors to estimate associated time courses for each subject: since the eigenconnectivities

are orthonormal we can reformulate the equation above as $\mathbf{C}_s - \overline{\mathbf{C}}_s = \mathbf{U}_{\text{red}}\mathbf{W}_s$, where $\mathbf{U}_{\text{red}}$ acts as the regressor. Thus, for a single set of "group" eigenconnectivities, which we can interpret as building blocks of dFC, we obtain subject-specific time-dependent weights $\mathbf{W}_s$ that describe their contribution in representing dFC matrix over time. We inspected the time-dependent weights to make sure that the eigenconnectivities contributed to dynamic connectivity in all subjects (i.e., non-zero weights for all subjects).

Comparing HC subjects and RRMS patients

We first compared the average dFC strength $\overline{\mathbf{C}}_s$ between HC subjects and RRMS patients using two-sample t tests. Statistically significant differences were determined using non-parametric permutation testing, by estimating a maximal statistic under the null hypothesis of no difference; i.e., observed differences in the 95^{th} percentile of the null distribution of differences between subjects randomly assigned to the two groups were significant at the 5% level ($n = 1'000$ randomizations, corrected p-value) (Nichols & Holmes, 2002).

We next assessed whether the contribution of the eigenconnectivities was different between HC subjects and RRMS patients by comparing the percentage of time-dependent weights with positive signs. A difference between groups would indicate that the contribution of a specific FC pattern is biased in the RRMS patients. We compared such a summary measure since individual time points are not consistent across subjects in RS studies. We tested for differences in the percentages using Hotelling's T^2 multivariate test. Hotelling's T^2 test can also be seen as a maximum univariate t^2 test for a specific linear combination of the means of the two groups; i.e., the weighted sum $\sum_{k=1}^{K} a_k(m_k^{HC} - m_k^{MS})$ is maximal, where $\sum_{k=1}^{10} a_k^2 = 1$, m_k is the average percentage for eigenconnectivity k across each group of subjects and HC and MS indicate the HC and RRMS subjects, respectively. Hotelling's T^2 test identifies group and interaction (group × eigenconnectivity) effects. For the interaction effect, the weighted sum is modified as $\sum_{k=1}^{K-1} a_k \{(m_k^{HC} - m_{k+1}^{HC}) - (m_k^{MS} - m_{k+1}^{MS})\}$. We estimated the effect size from the Mahalanobis distance (a multivariate generalization of Cohen's d), which is defined as: $D = \sqrt{T^2(1/n_1 + 1/n_2)}$, where $T^2 = FK(n_1 + n_2 - 2)/(n_1 + n_2 - K - 1)$ and n_1, n_2 are the sizes of the HC and RRMS groups and $K = 10$ is the number of variables (eigenconnectivities).

We calculated a summary measure of percentage positive weights according to the weights of the Hotelling T^2 test for each subject and assessed whether it was correlated with clinical disability (EDSS) using Pearson correlation within RRMS patients.

Phase randomized data

Handwerker et al. (2012) showed that the sliding-window technique can result in spurious dFC changes and it is thus crucial to compare the results to an appropriate null model. We phase randomized both brain *activity* time courses before estimating dFC (Handwerker *et al.*, 2012), and *connectivity* time courses estimated from original activity (Allen *et al.*, 2014) in separate simulations. Specifically, we Fourier transformed each subject's activity or connectivity time courses, randomized the phases independently for each brain region or connectivity pair, and transformed back to the time domain using the amplitude of the real data but randomized phases (Theiler *et al.*, 1992). Phase randomizations preserve the mean, variance and autocorrelation properties of the time courses, while randomizing their precise timing. Fluctuations in FC estimated from phase randomized brain *activity* are thus due to random activity timing. Randomized *connectivity* destroys its precise timing, while preserving

inter-subject variability in the mean and variance of connectivity (i.e., we can disentangle effects of inter-subject variability in mean FC values from FC fluctuations).

We also assessed the significance of each eigenconnectivity by comparing its eigenvalue to a null distribution obtained using the phase randomized data. Analytic methods to estimate the dimensionality currently employed for *activity* time courses, such as Laplace PCA (Minka, 2000), overestimate the dimensionality in the presence of correlated noise (Cordes & Nandy, 2006). Comparing the eigenvalues to a null distribution obtained using surrogate data avoids making parametric assumption about the noise. This approach is similar to how significant latent dimensions are estimated in PLS (McIntosh *et al.*, 1996; Krishnan *et al.*, 2011) and in other fields studying time series (e.g. Bjornsson & Venegas, 1997). The intersection of the real eigenvalue spectrum and the 95^{th} percentile of the phase randomized spectrum estimates the dimensionality of the data ($n = 100$ randomizations, maximal statistic to correct the p-value).

Motion

As described above, we scrubbed the data and regressed out head motion parameters from the regional time series. However, residual motion may still drive changes in dFC. We thus tested for relationships between the time-dependent weights and four summary measures of motion using Pearson's correlation. We converted FD defined above and root mean square displacement $D = \sqrt{\Delta x^2 + \Delta y^2 + \Delta z^2}$ to a measure for each window by estimating the intensity (average) and fluctuation of motion (standard deviation) in each window.

Statistically significant correlations with motion were determined using non-parametric testing, similar to above; i.e., observed correlations in the 95^{th} percentile of the null distribution of correlation with phase-randomized motion time series were significant at the 5% level ($n = 1'000$ randomizations, corrected p-value). Note that these phase randomizations were not the same as those used above to obtain a null distribution of dFC.

3.1.3 Results

dFC estimation

After scrubbing, each subject had between 401 and 440 volumes left, from which we estimated dFC matrices $\mathbf{C}_s$. For a window length of 30 TRs, the multi-subject data matrix $\mathbf{X}$ was of size $3'828 \times 5'554$, i.e., it contained $5'554$ unfolded correlation matrices across all subjects. Each subject had an average of $\overline{T} = 198 \pm 5$ windows and this number did not differ between HC subjects and RRMS patients (two-sample t test, $p = 0.8$). Fig. 3.4 and Supplementary Movie 1 show an example of the estimated dFC for a HC subject. FC is visibly dynamic over the course of the scan as indicated by the change of color across windows. For this subject, the majority of correlations were positive, but we also observed alternations between positive and negative correlations.

Eigenconnectivities

The average dFC strength $\overline{\mathbf{C}}_s$ across each group of subjects showed functional clusters typically observed in stationary FC analyses (Supplementary Fig. A.1): strongly connected visual areas (bilateral calcarine, cuneus, lingual gyri and other occipital regions), motor areas (bilateral pre- and postcentral gyri,

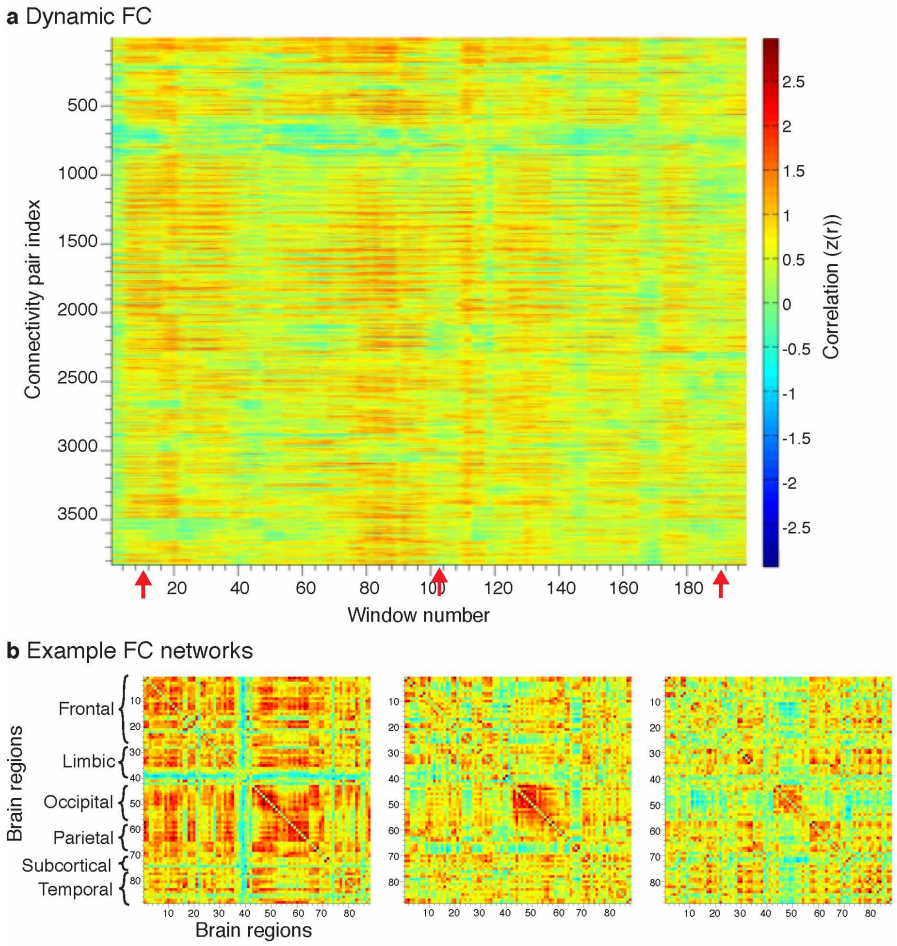


Figure 3.4 – (a) Example of dFC for a HC subject (unfolded and temporally concatenated, Fisher-z transformed correlation matrices). (b) Example FC networks, corresponding to the windows marked with red arrows in (a) and reshaped into correlation matrices.

rolandic opercula, supplementary area, paracentral lobule), and regions of the DMN (bilateral superior frontal gyri, cingulate gyri, angular gyri and precuneus).

We then identified dFC patterns across all HC subjects and RRMS patients by calculating the eigenvectors of the covariance matrix of the concatenated, demeaned dFC matrices. The first ten eigenconnectivities retained 34% of the variance or equally, approximated the data with an accuracy of 34% (Supplementary Fig. A.2 shows all 3828 eigenvalues). The intersection of the real and phase-randomized eigenvalue spectra estimated that the first 30 eigenconnectivities predominantly carried signal (5% corrected significance level, Supplementary Fig. A.2). In the remainder of this manuscript we restrict our analysis to the first 10 eigenconnectivities, which have eigenvalues well above the significance level, and because the multivariate Hotelling's T^2 test sets an upper limit on the number of components due to group sizes.

Fig. 3.5 shows some of the estimated “group" eigenconnectivities, which represent orthogonal (independent) connectivity patterns with coherent fluctuations. Since the mean FC strength of each pair has been removed for each subject, these patterns encode how dFC fluctuates around the average connectivity strength over time. The first few eigenconnectivities, which explain most of the variance

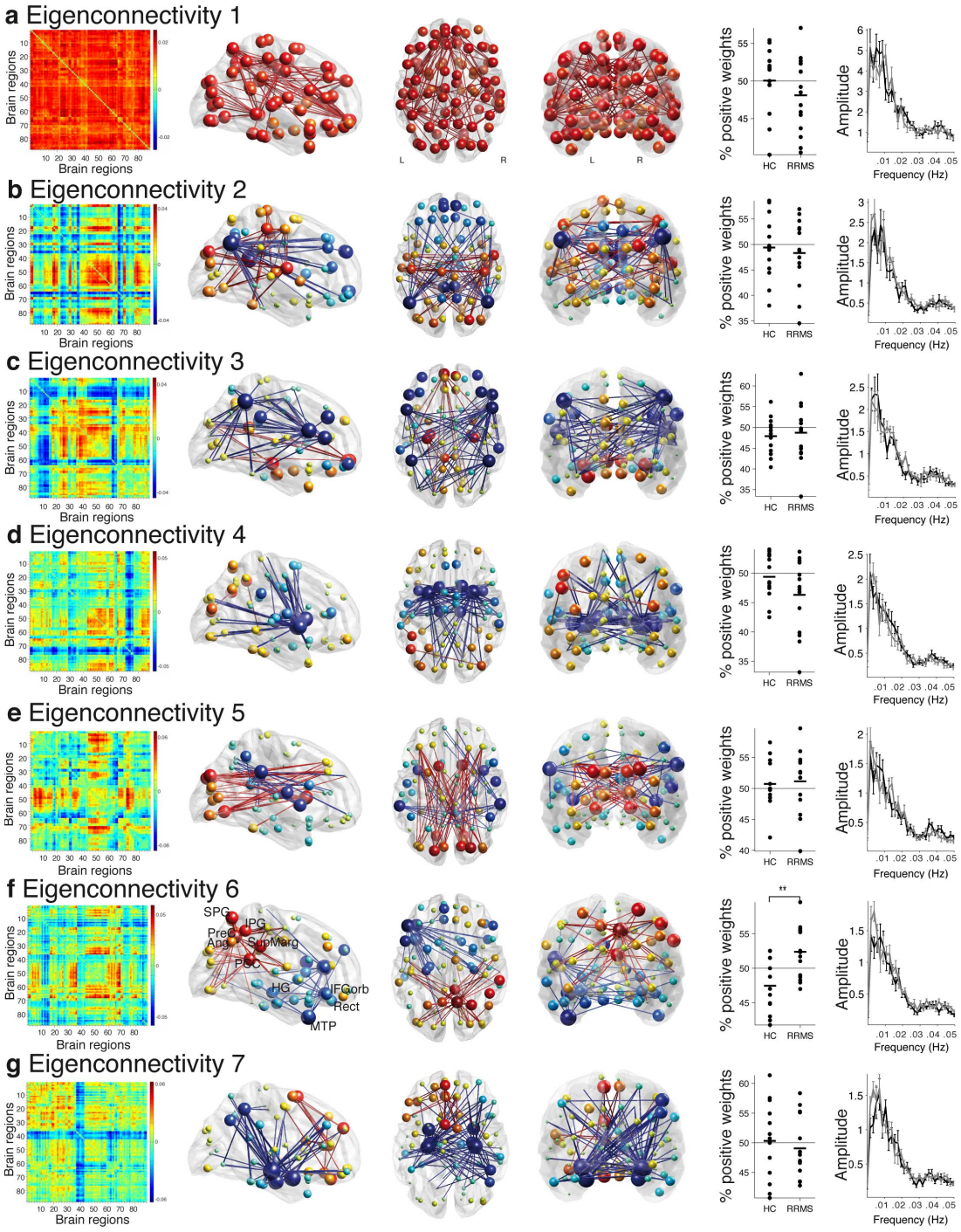


Figure 3.5 – Continued on next page

in the dFC matrices, represent distributed connectivity patterns. Eigenconnectivity 1 identifies an all-positive pattern and reflects overall FC strength (“excursions" around the mean where whole-brain dFC is globally increased or decreased; Fig. 3.5a): connections between superior frontal regions (23 – 26)

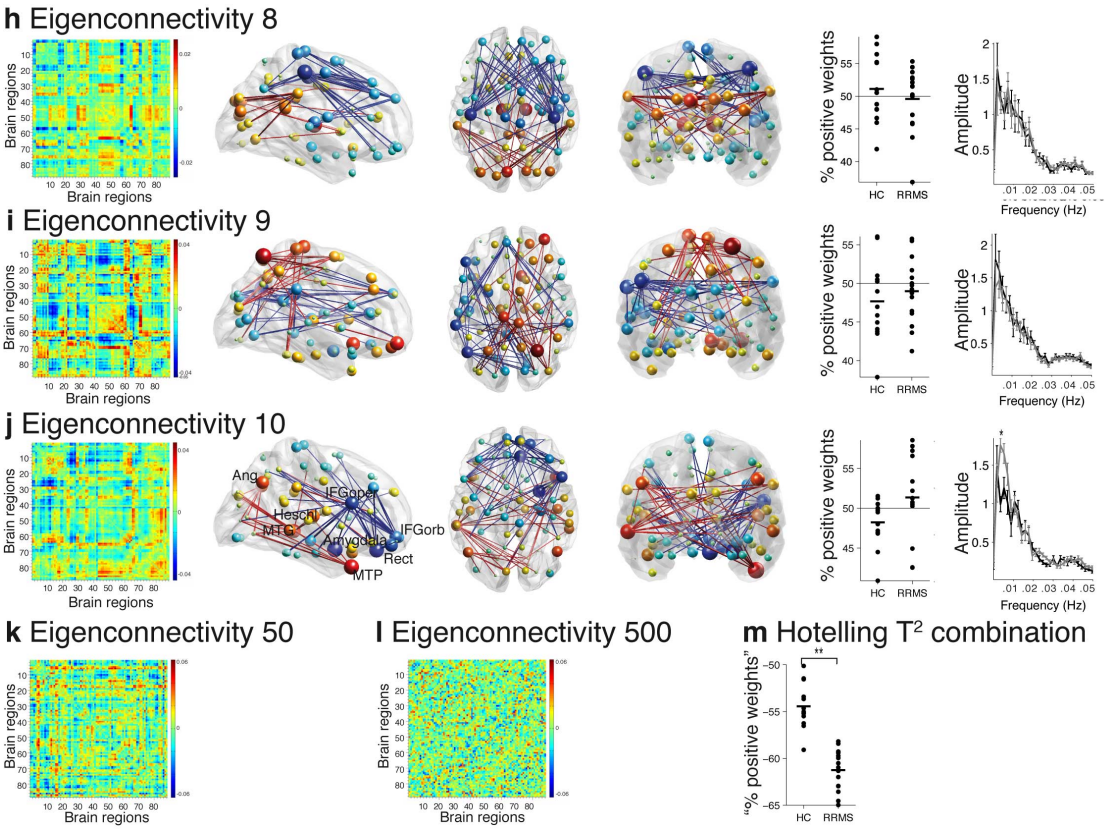


Figure 3.5 – Eigenconnectivities estimated from dFC across all 28 subjects show large-scale connectivity patterns. (a-j) Eigenconnectivities 1-10 in matrix representation (column 1), and sagittal, axial and coronal views of the 2% strongest connections in brain space (columns 2-4). Brain regions are shown as spheres where their size represents their degree and the color their sign (red for positive and blue for negative degrees). Connections use the same color convention (red = positive, blue = negative). Percentage of positive weights of each eigenconnectivity over time (column 5) and amplitude spectra of the time-dependent weights for HC subjects (black) and RRMS patients (gray), error bars represent standard errors (column 6). (k, l) Eigenconnectivities 50 and 500, (m) linear combination of the percentages according to the weights of Hotelling's T^2 test; combined percentage per subject (circle), group means for HC subjects and RRMS patients, respectively (line). Ang, angular gyrus; HG, hippocampal gyrus; orb, orbital; oper, operculum; Rect, gyrus rectus; SupMarg, supramarginal gyrus

and other regions in the brain were strong (dark red areas) while connections of limbic (38 – 42) and subcortical regions (71 – 74) were weaker (yellow areas).

The following eigenconnectivities have both positive and negative patterns (red and blue areas in Fig. 3.5, respectively). Windows with strong FC in red areas and weak FC in blue areas will have large positive weights for a given eigenconnectivity. Windows with weak FC in red areas and strong FC in blue areas on the other hand will have large negative weights. Therefore, an eigenconnectivity contrasts windows with strong FC in red areas and weak FC in blue areas with those of the opposite character. Also, we can interpret the connectivity patterns as representing connections that covary (i.e., within red areas and within blue areas) or anti-covary (i.e., between red and blue areas) according to that dimension of the representation. It should be noted that the sign (i.e., positive and negative for red and blue, respectively) is arbitrary and should not be confounded with anti-correlation for negative signs.

Eigenconnectivity 2 contrasts a network of connections involving mostly regions of the DMN (blue)

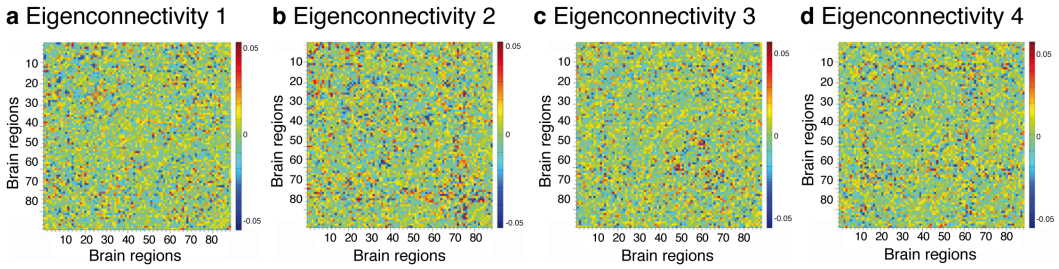


Figure 3.6 – Eigenconnectivities 1-4 estimated from phase-randomized dFC matrices lose all structure.

with regions involved in motion, vision, speech and hearing (red), i.e., it contrasts windows with strong FC in DMN regions with windows with strong FC in sensori-motor regions (Fig. 3.5b): Blue areas include connections of medial frontal (23 – 26), anterior (ACC) and posterior cingulate gyri (PCC; 31,32,35,36), angular gyri (65,66) and precuneus (67,68). Red areas include connections of primary sensori-motor areas (1,2,18 – 20,57,58,69,70), occipital, ventral and dorsal visual areas (43 – 56,59,60), speech (17,63,64,79,80), auditory areas (77 – 82), and the insula (29,30). Eigenconnectivity 3 identifies a network between regions involved in memory and higher cognitive processing (25 – 28,35,37 – 39,67,68), the rolandic operculi (17,18) and Heschl's gyri (77,78; red), and another network of connections of the inferior (IFC) and middle frontal cortices (7 – 14), and inferior parietal cortices (IPC; 61,62; blue; Fig. 3.5c). Eigenconnectivity 4 identifies a network of connections for thalamic basal ganglia (71 – 76) and the insula (29,30; blue), and another network of regions involved in visuo-spatial processing (43 – 56,59 – 62,83 – 88; red, Fig. 3.5d).

Fig. 3.5 also shows that most FC patterns contain connections of regions situated in both hemispheres, with the exception of eigenconnectivity 6 where the network of blue connections lies predominantly in the left hemisphere (inferior frontal areas) and red ones in the right hemisphere (parietal areas); and eigenconnectivity 9 where most blue connections lie in the left hemisphere.

Eigenconnectivities of higher order, such as e.g. eigenconnectivity 500, resemble purely artefactual components that do not identify meaningful functional clusters (Fig. 3.5l).

Results with phase-randomized activity or dFC resulted in eigenconnectivities with no apparent structure (Fig. 3.6, Supplementary Fig. A.4e) and little explained variance (5% for the first ten eigenconnectivities for both randomizations).

Time-dependent contribution of eigenconnectivities

Fig. 3.7a shows the time-dependent weights $\mathbf{W}_s$ obtained for the same example subject as in Fig. 3.4 by projecting the dFC matrix onto the first ten eigenconnectivities (see Supplementary Fig. A.3 for all subjects). The weights varied smoothly across time but presented pronounced variability, again indicating that FC is highly dynamic. The 5^{th} and 95^{th} percentiles of weights observed for phase-randomized correlation time courses indicate that the observed weights were both within and outside of ranges observed for the surrogate data, and for which time points there is strong evidence for a large contribution of an eigenconnectivity.

The weights of eigenconnectivity 1 (dark blue) track the overall change in FC strength and perfectly correlate with the mean correlation of each window (column-sum of $\mathbf{C}_s$), $\rho = 1, p < 0.001$ for all subjects: at time points when whole-brain dFC is strong, or in other words above average dFC strength (red areas

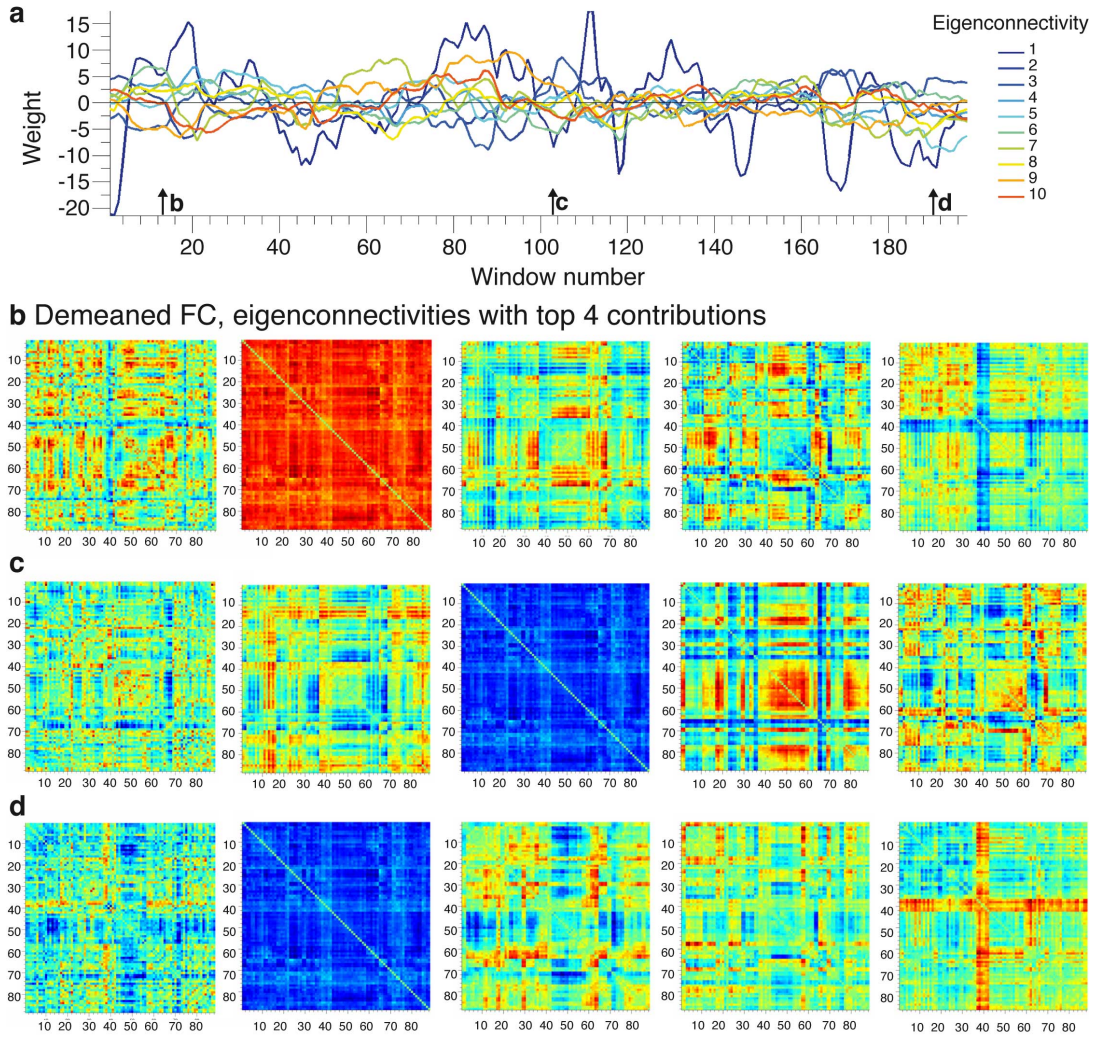


Figure 3.7 – (a) Time-dependent weights of the 10 first eigenconnectivities for the same example subject as in Fig. 3.4. Straight lines represent the 5th and 95th percentiles of phase-randomized weights across all subjects and eigenconnectivities. (b-d) Example FC networks of Fig. 3.4c after row-wise demeaning (first column) and the eigenconnectivities with the four largest weights (2nd to 5th column). We flipped the sign of eigenconnectivities with negative weights (e.g. in (c) eigenconnectivity 1 had a negative weight and thus appears as “blue"). (b), (c) and (d) correspond to the windows marked with arrows in (a).

in Fig. 3.4a), the weight is positive, while at time points when whole-brain dFC is weak (blue areas in Fig. 3.4a), the weight is negative. The weights of the other eigenconnectivities track more subtle changes in dFC patterns.

Figs. 3.7b-c show which four eigenconnectivities contributed most strongly to the demeaned example dFC networks of Fig. 3.4c. The eigenconnectivities with strongest contributions capture different aspects at a specific window, such as the strong positive occipito-parietal FC and weak FC between limbic regions and the rest of the brain in Fig. 3.7b.

The weights were highly variable for all subjects, which indicated that the eigenconnectivities captured inherent structure in the fluctuations across all subjects and not just within a single subject.

Comparing HC subjects and RRMS patients

First, we examined whether the average dFC strength $\overline{\mathbf{C}}_s$ differed between HC subjects and RRMS patients. Supplementary Fig. A.1 indicates a widespread reduction in FC strength in RRMS patients and two-sample t tests revealed 22 significantly weaker connections, mainly between the left amygdala and occipital and parietal regions; and connections of the middle and posterior cingulate gyri and bilateral superior frontal gyri ($p < 0.05$, corrected, see Supplementary Table A.2 for a full list).

Next, we compared the percentage of positive weights of the ten first eigenconnectivities between HC subjects and RRMS patients (Fig. 3.5, column 5). HC subjects and RRMS patients differed strongly in the percentage of positive weights, indicating a bias in the contribution of some eigenconnectivities ($F_{(10,17)} = 4.1,\ p = 0.005,\ D = 3.0$ for the group effect; $F_{(10,17)} = 3.5,\ p = 0.01,\ D = 2.8$ for the interaction effect; Fig. 3.5m). The Hotelling T^2 group difference was $\sum_{k=1}^{K} a_k(m_k^{HC} - m_k^{MS}) = 6.8 \pm 5.3\%$, with $a_1 = 0.06, a_2 = 0.13, a_3 = -0.09, a_4 = 0.17, a_5 = -0.11, a_6 = -0.72, a_7 = 0.09, a_8 = 0.10, a_9 = -0.18, a_{10} = -0.63$. The most important differences thus stemmed from eigenconnectivities 6 and 10, which was confirmed in post hoc two-sample t tests ($p = 0.002, d = -1.29$; $p = 0.06, d = -0.74$, respectively, $p > 0.1$ for the other 8 eigenconnectivities). For both eigenconnectivities RRMS patients had more positive contributions than HC subjects (mean percentage > 50%), i.e., their red connections were more often strongly connected (positive weights increase the FC of red areas, and "strongly" here means above average) and their blue ones were more often weakly connected (positive weights reduce the FC of blue areas). Red connections of eigenconnectivity 6 were for example concentrated in posterior regions of the DMN (superior parietal gyrus, precuneus, angular gyri, suparmarginal gyri), while blue connections lay mostly in anterior (IFC, gyri recti) and midtemporal regions (Fig. 3.5f). Red connections of eigenconnectivity 10 included temporal regions (middle temporal cortex and pole, Heschl), and the angular gyri, while blue connections spanned a network of orbito-frontal regions, the right amygdala and motor regions (precentral gyrus, paracentral lobule; Fig. 3.5j). Fig. 3.5j also shows that many connections in eigenconnectivity 10 were across hemispheres and ran along the anterior-posterior axis. To explore whether the percentage positive weights was dominated by small weights, we repeated the analysis after excluding the smallest 10% and 20% weights and obtained similar results ($F_{(10,17)} = 4.2,\ p = 0.005,\ D = 3.1$ for the group effect, post hoc tests significant for the sixth $p = 0.006$ and tenth components $p = 0.038$; and $F_{(10,17)} = 3.8,\ p = 0.008,\ D = 2.9$, post hoc tests significant for the sixth $p = 0.02$, and 10th components $p = 0.02$, respectively). The percentage positive weights of the first ten eigenconnectivities for phase-randomized data were not different between HC subjects and RRMS patients (Hotelling T^2, $p > 0.4$ for both group and interaction effects for both randomizations). The combination of unstructured eigenconnectivities and altered dynamics suggests real fluctuations drove the results obtained from the original data. The Hotelling T^2 combination of percentage positive weights did not correlate with EDSS.

The time-dependent weights of eigenconnectivity 10 showed a significantly increased frequency amplitude at 0.003 Hz in RRMS patients ($d = -1.5$, $p < 0.02$, corrected, Fig. 3.5).

We then varied the window length from short (30 seconds) to long (2 minutes) and increased step sizes from 2 to 8 TRs. Supplementary Figs. A.4a-d show that while the first eigenconnectivities estimated from different window lengths roughly resemble each other, the estimated connectivity patterns are not the same. Notably, the covarying FC between middle and inferior frontal regions apparent in eigenconnectivity 3 at a window length of 30 TRs is not detected at longer windows lengths. We also observe that eigenconnectivity 5 breaks into two patterns, 4 and 6, at a window length of 120 TRs.

In general, the visible block structure of eigenconnectivities progressively vanishes at larger window lengths. The Hotelling T^2 test did not reveal significant differences between the average weights of HC subjects and RRMS patients for window lengths above 30 TRs ($p > 0.2$).

In exploratory work, we also repeated our analysis using different band pass filters and obtained similar results in terms of eigenconnectivities and explained variance for a high-frequency cut-off of 0.2 Hz and 0.3 Hz, but no significant group difference ($p = 0.12, 0.08; 0.64, 0.75$ for group and interaction effects and the two cut-offs respectively).

Influence of intra-individual differences

We then repeated our analysis without row-wise demeaning each subject's dFC matrix but only subtracting the average dFC across time and subjects. The first ten eigenconnectivities explained 42% of the variance in the data. They resembled the FC patterns estimated from the subject-wise demeaned data but were not the same (Fig. 3.8a). For example, the first eigenconnectivity shows much more pronounced differences between limbic and subcortical connections (more yellow than in Fig. 3.5), the blue connections of the 2nd one bear little resemblance, within frontal and within visual connections are less pronounced for the 3rd one, and the 6th one bears little resemblance. Eigenconnectivity 6 includes connections of parietal regions ($59-62$), and the thalami ($75, 76$) with the rest of the brain, connections with visual regions (red); and the putamen ($73, 74$) to the rest of the brain and between motor ($1, 2, 19, 20, 57, 58$) and frontal regions (blue).

The contribution of the first ten eigenconnectivities was significantly different between the two groups (group effect $F_{(10,17)} = 2.7, p = 0.04,\ D = 2.4$, interaction effect $F_{(10,17)} = 3.0, p = 0.02, D = 2.6$). The Hotelling T^2 group difference was $\sum_{k=1}^{K} a_k(m_k^{HC} - m_k^{MS}) = -48 \pm 46\%$, with $a_1 = -0.76, a_2 = 0.02, a_3 = 0.03, a_4 = -0.20, a_5 = 0.04, a_6 = -0.01, a_7 = 0.30, a_8 = -0.32, a_9 = 0.43, a_{10} = -0.03$. The largest contributions thus came from the first and seventh to ninth eigenconnectivities. The percentage of positive weights for the first eigenconnectivity were $66 \pm 27\%$ for HC subjects and $29 \pm 22\%$ for RRMS patients ($p < 0.001$). As this is an all-positive pattern, this again indicates that RRMS patients had a reduced overall FC strength. Post hoc t tests also indicated a trend for eigenconnectivity 7 ($61 \pm 32\%, 40 \pm 23\%, p = 0.052$). As we did not demean each subject's dFC matrix, large positive percentages can here be interpreted as reflecting stronger FC in red areas, and low percentages as reflecting stronger FC in blue areas (positive weights increase the FC of red areas, negative weights that of blue areas). Eigenconnectivity 7 contrasted connections centered on the right PCC and left Heschl gyrus (red), with connections of limbic regions (hippocampi (HP), PAH, amygdala), the left thalamus and the left middle temporal pole (blue; Fig. 3.8b). The Hotelling T^2 combination of percentage positive weights did not correlate with EDSS.

Eigenconnectivities estimated from phase-randomized dFC matrices showed structured patterns (Fig. 3.8c), still retained a large amount of variance (23% for the first ten), and the average weights differed significantly between the groups ($F_{(10,17)} = 3.4,\ p = 0.01,\ D = 2.7$ for the group effect, $F_{(10,17)} = 2.6,\ p = 0.04,\ D = 2.4$ for the interaction effect). These results indicate that average correlation values rather than FC fluctuations drove the PCA and the group difference if no individual demeaning is performed.

Eigenconnectivities estimated from phase-randomized activity time courses showed no structure (Supplementary Fig. A.4f), explained 5% of variance for the first ten and Hotelling's T^2 was not significant as this simulation did not preserve the mean correlation values.

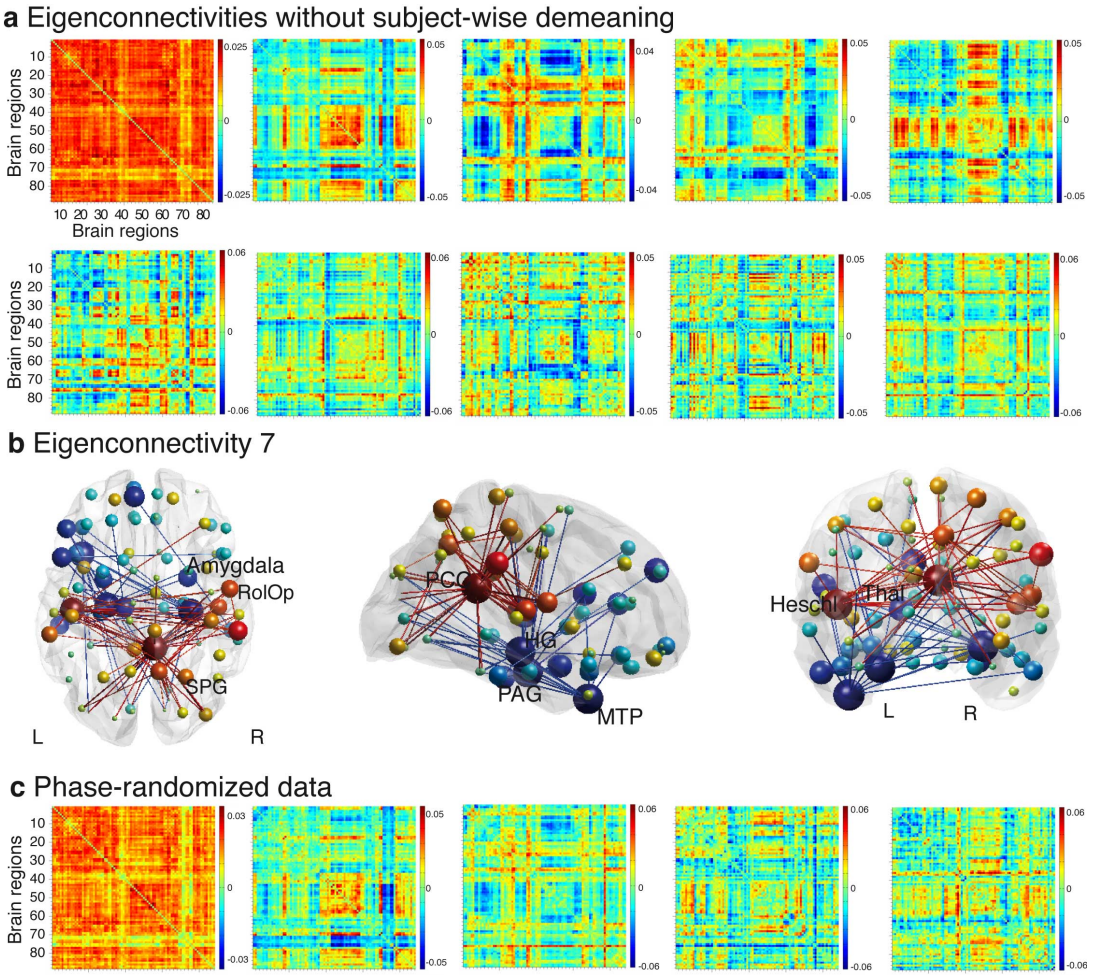


Figure 3.8 – (a) Eigenconnectivities 1-6 estimated from the concatenated dFC matrices without subject-wise demeaning. (b) Axial, sagittal and coronal views of the strongest connections of eigenconnectivity 7. Connections in red were temporally overrepresented in HC subjects, and blue ones in RRMS patients. Only the 2.5% strongest connections are shown for visualization purposes. (c) Eigenconnectivities estimated from phase-randomized dFC matrices retain large-scale structure.

Motion

Motion parameters for each subject were small (Supplementary Table A.3). The four summary motion measures were not significantly different between the two groups (HC subjects and RRMS patients respectively, FD: mean 0.11, 0.11; SD 0.05, 0.05; D: mean 0.05, 0.05, SD 0.03, 0.03, $p > 0.4$ for all). Dynamic weights of the eigenconnectivities were not significantly correlated with any of the four motion parameters, except for two HC subjects (time-dependent weights of eigenconnectivity 4 with the standard deviation of FD, $R^2 = 0.64$; time-dependent weights of eigenconnectivity 9 with the mean and standard deviation of D, $R^2 = 0.75, 0.73$ respectively).

As scrubbing may affect the estimation of dFC, we repeated the analyses without scrubbing and found highly similar results in terms of explained variance, extracted connectivity patterns and altered contributions in MS patients.

3.1.4 Discussion

In this study, we introduce a novel data-driven approach for extracting dFC patterns in the resting brain. This exploratory technique applies PCA to dynamic whole-brain FC of multiple subjects during resting state. PCA identified FC patterns, the eigenconnectivities, that capture FC pairs with similar dynamics. We found large-scale connectivity patterns that encompassed regions from different functional clusters. The temporal contributions of these eigenconnectivities can be obtained by projection and reveal their importance in explaining dFC over time. On the example of MS, we identified a subset of connections with altered dynamics in minimally-disabled RRMS patients.

Related methodological studies

Previous studies reported large-scale reorganizations of *whole-brain* functional brain networks during learning (Bassett *et al.*, 2011), task preparation (Ekman *et al.*, 2012), or rest (Allen *et al.*, 2014). Bassett et al. (2011) studied dynamic community organization, Ekman et al. (2012) predicted upcoming tasks and errors from network measures (e.g. degree), and Allen et al. (2014) employed clustering to identify stable FC topologies. Our approach employs PCA to characterize the most abundant building blocks of dFC and, like previous studies, identifies prominent changes in network topology. In contrast to "fixed" FC states detected using clustering, the eigenconnectivities overlap in time and their combined contributions give rise to varying network topologies. As such, they represent basic building blocks of network configurations that occur over time, rather than discrete states. However, clustering could also be applied to the time-dependent weights to identify eigenconnectivities with similar temporal profiles (in contrast to Allen et al. (2014) who clustered windowed correlations). We here opted for another approach and looked at summary measures of the contributions of these patterns over time.

The most prominent technique to study spontaneous brain *activity* and to identify voxels with similar activity is spatial ICA. To identify *connections* with similar fluctuations, we used PCA, which optimizes for explained variance and orthogonality, and gives an efficient, linear representation. One could also consider ICA as an alternative decomposition to extract connectivity patterns from multisubject dFC matrices. It is however not clear if the assumption of statistical independence would lead to better components and, because of the independence assumption, ICA components have no inherent ordering, which would complicate the determination of relevant ones. Smith et al. (2012) recently applied a cascade of spatial and temporal ICA to fast-TR data to detect functionally (temporally) distinct, and spatially overlapping networks. While both studies are interested in the architecture of the brain during rest, the authors decomposed brain activity, while we studied brain connectivity. As such, they identified networks of brain regions with similar activity, whereas we identified networks of connections with similar fluctuations. Temporal ICA might not be appropriate to decompose dFC because dFC fluctuates slowly due to the sliding window estimation (temporal ICA needs a large number of time points to perform robustly) and because temporal ICA implicitly assumes temporal stationarity, i.e., constant moments across time (see also Calhoun *et al.*, 2012; Allen *et al.*, 2014).

Stationary and dFC in RRMS patients

We identified eigenconnectivities 6 and 10 that contributed differently between HC subjects and RRMS patients, indicating aberrant dFC in RRMS patients (Figs. 3.5f, j): The affected connections were concentrated in parietal regions such as the PCC, superior parietal gyrus and angular gyrus (more

frequently strongly connected in RRMS patients), and prefrontal regions (e.g. orbitofrontal gyrus, gyrus rectus) and the amygdala (more frequently weakly connected). The importance of these regions suggests that the dFC of at least part of the DMN is altered in patients. These results add to, and complement, several studies which have reported disturbed stationary FC in different parts of the DMN in MS patients (Bonavita *et al.*, 2011; Hawellek *et al.*, 2011; Richiardi *et al.*, 2012; Rocca *et al.*, 2012). The distinction between posterior and anterior/temporal DMN regions seen in eigenconnectivity 6 is similar to the DMN fragmentation observed in Allen et al. (2014) and would be missed in conventional stationary FC analyses.

When repeating the analysis without removing each subject's average connectivity strength, we found distinct patterns of connections that differed between HC subjects and RRMS patients, which were average- rather than dynamically-driven (Fig. 3.8). The identified brain regions are largely consistent with those in a previous study by our group, which aimed at discriminating between HC subjects and RRMS patients based on stationary FC (Richiardi *et al.*, 2012); i.e., the left Heschl's gyrus, right rolandic operculum, and right superior parietal gyrus, were identified as being more strongly connected in HC subjects in both studies. Reduced stationary FC in MS patients is indeed commonly observed and is thought to stem from structural damage (Lowe *et al.*, 2002; Roosendaal *et al.*, 2010; Bonavita *et al.*, 2011; Richiardi *et al.*, 2012). However, we also confirmed the bilateral parahippocampal gyri, right amygdala, left thalamus and left midtemporal pole as more strongly connected in RRMS patients (Richiardi *et al.*, 2012). Such selective increased FC has been reported by several other studies as well (Roosendaal *et al.*, 2010; Bonavita *et al.*, 2011; Hawellek *et al.*, 2011; Richiardi *et al.*, 2012; Rocca *et al.*, 2012; Dogonowski *et al.*, 2013) and has been interpreted as a compensatory mechanism (Roosendaal *et al.*, 2010; Bonavita *et al.*, 2011) because activation studies have shown increased activity in regions devoted to the task and the recruitment of additional regions in MS patients (Pantano *et al.*, 2002; Rocca *et al.*, 2005). It is however also possible that these increases originate from a loss of flexibility in functional interactions, rather than as a compensatory mechanism (Hawellek *et al.*, 2011). Interestingly, the connections in eigenconnectivity 10 obtained from the dFC analysis showed an increased low-frequency amplitude in RRMS patients, which might support the latter hypothesis.

While we did not observe correlations between the Hotelling T^2 combinations of positive percentages and clinical disability for the dynamically- or average-driven analysis, it should be noted that our sample size was relatively small (15 patients), and our patients minimally disabled (EDSS range $1.5-2.5$, meaning that they can live a normal life). It would be interesting to study changes in dFC in larger samples of MS patients, at different phases of the disease, and using additional clinical measures. Nonetheless, given the complementary insights gained from the analysis of dFC, such as which connections fluctuate coherently or which connections fluctuate more slowly, we believe that studies of dFC can help in understanding the functional impact of MS. However, we note that the choice of sliding window length and band pass filter affected our results as opposed to scrubbing or the removal of small weights. Therefore, while our results are promising, studies in independent cohorts are needed to further explore these effects and validate our findings.

Timescales and potential confounds of dFC

Correlations estimated from short windows have more variance due to the small number of samples but are more sensitive to changes in FC since long windows reduce temporal variability and may miss anticorrelations (Chang & Glover, 2010; Hutchison *et al.*, 2013b). Some studies have reported that 30

seconds of data suffices to discriminate between cognitive states and to estimate reliable modular graph metrics (Shirer *et al.*, 2012; Jones *et al.*, 2012). We therefore chose a sliding window length of 30 TRs to capture changes in FC over time. Given the similarities between estimated eigenconnectivities at different window lengths, it seems unlikely that it is random variability that drove FC patterns at short window lengths. Exploratory work using different band pass cut-off frequencies also suggested that the large-scale structure of the eigenconnectivities was largely preserved across broad frequency ranges. In our main analysis we chose a cut-off of 0.15 Hz as sliding window correlations are estimated from a small number of samples and a lower cut-off helps to remove noise. The chosen cut-off is similar to what has been used in other studies on dFC (Bassett *et al.*, 2011; Kang *et al.*, 2011; Handwerker *et al.*, 2012; Allen *et al.*, 2014; Li *et al.*, 2014), but it would be interesting to further investigate the effects of different frequency cut-offs on dFC when physiological data is available to reduce noise in resting-state activity time courses. Together, these results suggest that patterns of covarying connections were largely similar across varying window lengths, frequency cut-offs and with or without scrubbing.

We also tested for correlations between the contributions of the eigenconnectivites across time and motion measures. Out of the 28 subjects, these contributions were only significantly correlated for two subjects and one of the eigenconnectivities. While we cannot exclude that motion-related or physiological noise is potentially driving the observed FC patterns, the large-scale organization of these patterns, the importance of the precise timing of FC fluctuations as shown by the results with phase-randomized dFC and the minimal correlations with motion parameters, suggest that these patterns are not simply the result of noise.

Chang et al. (2013b) recently linked heart rate variability to dFC in regions associated with vigilance and arousal, suggesting the contribution of autonomic factors to the observed nonstationarity. Chang and colleagues also showed that alpha band power in EEG was related to FC fluctuations between the DMN and the dorsal attention network (Chang *et al.*, 2013a). Further concurrent EEG-fMRI studies will be crucial in clarifying the relationship between these fluctuations and neural dynamics, as well as how they relate to attentional states. Along those lines, the scale-free dynamics of EEG microstate sequences have been shown to reach timescales of fMRI resting-state fluctuations (Van de Ville *et al.*, 2010) and the link between EEG microstates and dynamic fMRI FC is an intriguing future research question.

Variability as a signal in its own right

Not only has interest in the variability of brain *connectivity* increased, but several recent studies have also highlighted that the variability of brain *activity* is a signal of interest and provides information complementary to the analysis of mean brain activity. Notably, brain activity variability has been linked to task performance, development and disease and it has been suggested that signal variability is not “noise" but crucial for spontaneous or stimulated reorganizations of the brain (McIntosh *et al.*, 2008; Garrett *et al.*, 2010). Just as studies on brain *signal* variability provided novel insights into brain function, we believe that studies on brain *connectivity* variability will further our understanding of the brain’s functional organization. Calhoun et al. (2012) and Starck et al. (2012) recently reported changes in dFC variability in patients with schizophrenia and autism. We believe that our proposed approach can go beyond identifying changes in dFC variability and reveal patterns in FC fluctuations. Our approach is also applicable to other neurodegenerative diseases and the temporal evolution of the brain’s functional network at longer time scales, such as during learning or development.

Methodological limitations and future directions

In this study we used anatomically defined regions and a simple linear estimator of dFC, but other parcellations or more complex estimators of dFC could easily be used with the proposed approach.

A comparison between the real eigenvalues and those obtained from the phase-randomized data suggested that the first 30 components predominantly carried signal, while the latter ones were not statistically different from noise. As the eigenvalues were close in magnitude, it seems however reasonable to not take this number too literally but rather as an indication of the overall range. An alternative method to estimate the dimensionality of our data, Laplace PCA (Minka, 2000), grossly overestimated its dimensionality (1362 dimensions). The comparison of the eigenvalues to those obtained from surrogate data on the other hand provided a reasonable, rough estimate to guide the choice of dimensionality. Other dimensionality estimation approaches, such as those used in activation studies (Yourganov *et al.*, 2011) or ICA (Varoquaux *et al.*, 2010), could be adapted to dFC and evaluated in future studies. The ten first eigenconnectivities used for further analysis did not use the full “signal space" and explained 34% of the variance in dFC, but our results suggest that they captured important differences between the groups. To identify connectivity patterns across subjects (rather than modelling subject-specific patterns or noise), it might indeed not be necessary to explain a majority of the variance. This is also highlighted by recent work of Li et al. (Li *et al.*, 2014) who identified 16 FC states using a dictionary learning approach and whose optimal model had residuals of around 50%.

It should also be noted that PCA is a linear decomposition and only sensitive to linear relations among connectivity pairs. To identify FC patterns based on nonlinear relations, one could for example employ kernel PCA, an extension of PCA, which has shown promising results for face recognition (Schölkopf *et al.*, 1998).

Finally, we also mention that other measures should be explored to exploit the full potential of temporal dynamics of eigenconnectivity time-dependent weights, in addition to the proposed percentage of positive weights.

In conclusion, we proposed a novel approach to estimate building blocks of dynamic brain connectivity and assessed their disturbance in minimally disabled RRMS patients.

3.2 FC patterns estimated by canonical correlation analysis

In the journal article presented in this chapter we simply concatenated the data from all subjects and then applied PCA to identify common dFC patterns across multiple subjects. Instead, we could search for FC patterns that take commonalities across datasets, but not aspects specific to individual subjects, into account. A classical method to identify such common patterns is canonical correlation analysis (CCA) (Hotelling, 1936), and its generalzation (gCCA) to more than two datasets (Kettenring, 1971). gCCA searches for a low-dimensional representation that takes statistical dependencies between datasets, as measured by their correlations, into account. If each dataset is whitened[2] in a first step, gCCA, however, simplifies to PCA. This property has for example been exploited in the analysis of gene expression (Tripathi *et al.*, 2008) or as a preprocessing tool to enhance the reproducibility of ICA (Varoquaux *et al.*, 2010). Specifically, Varoquaux *et al.* (2010) proposed a two-level model for multi-subject fMRI data, which is based on a subject-level PCA that reduces both the dimensionality of the data and subject-specific variability, that is followed by a group-level PCA (corresponding to a gCCA) capturing the remaining variability and identifying a common subspace.

In the next section we derive the equivalence of applying gCCA or the two steps of (i) whitening each subject's dFC and (ii) a group-level PCA. In the section thereafter we compare PCA and gCCA by applying it to the same dataset we used in the journal article.

3.2.1 gCCA and its relation to PCA

We first introduce CCA, which is defined for *two* datasets (Hotelling, 1936). For $S = 2$ datasets $\mathbf{X_1}$ ($C \times T_1$) and $\mathbf{X_2}$ ($C \times T_2$), which in our case are two subject's centered dFC matrices, CCA maximizes the correlation between the projections $\mathbf{X_1 A_1}$ and $\mathbf{X_2 A_2}$ onto a K-dimensional common subspace, with $\mathbf{A_1}$ ($T_1 \times K$), and $\mathbf{A_2}$ ($T_2 \times K$). CCA can be solved using a generalized eigenvalue decomposition:

$$\begin{bmatrix} 0 & \mathbf{C_{12}} \\ \mathbf{C_{21}} & 0 \end{bmatrix} \begin{bmatrix} \mathbf{A_1} \\ \mathbf{A_2} \end{bmatrix} = \Lambda \begin{bmatrix} \mathbf{C_{11}} & 0 \\ 0 & \mathbf{C_{22}} \end{bmatrix} \begin{bmatrix} \mathbf{A_1} \\ \mathbf{A_2} \end{bmatrix}, \tag{3.1}$$

where $\mathbf{C_{11}} = \mathbf{X_1}^T \mathbf{X_1}$, and $\mathbf{C_{22}} = \mathbf{X_2}^T \mathbf{X_2}$ are the auto-covariance matrices, $\mathbf{C_{12}} = \mathbf{X_1}^T \mathbf{X_2} = \mathbf{C_{21}^T}$ is the cross-covariance matrix and Λ contains the generalized eigenvalues on its diagonal.

For $S > 2$ datasets we use a generalization for which the solution is again given by a generalized eigenvalue decomposition (Kettenring, 1971; Varoquaux *et al.*, 2010; Afshin-Pour *et al.*, 2012):

$$\begin{bmatrix} 0 & \mathbf{C_{12}} & \mathbf{C_{13}} & \dots & \mathbf{C_{1S}} \\ \mathbf{C_{21}} & 0 & \mathbf{C_{23}} & \dots & \mathbf{C_{2S}} \\ \vdots & & \ddots & & \vdots \\ \mathbf{C_{S1}} & \dots & \dots & \dots & 0 \end{bmatrix} \begin{bmatrix} \mathbf{A_1} \\ \mathbf{A_2} \\ \vdots \\ \mathbf{A_S} \end{bmatrix} = \Lambda \begin{bmatrix} \mathbf{C_{11}} & 0 & \dots & 0 \\ 0 & \mathbf{C_{22}} & \dots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \dots & \dots & \mathbf{C_{SS}} \end{bmatrix} \begin{bmatrix} \mathbf{A_1} \\ \mathbf{A_2} \\ \vdots \\ \mathbf{A_S} \end{bmatrix}. \tag{3.2}$$

When all datasets are white, that is decorrelated and scaled to unit-variance ($\mathbf{X_s}^T \mathbf{X_s} = \mathbf{I}$ for all

[2] Whitening transforms a dataset $\mathbf{X}$ into a new dataset $\tilde{\mathbf{X}}$ whose covariance is the identity matrix, that is it is decorrelated and has unit variance. For row zero-mean data, the whitening transform can be found by solving the following eigenvalue decomposition: $\mathbf{X}^T\mathbf{X} = \mathbf{V}\Lambda\mathbf{V}^T$. Whitened data $\tilde{\mathbf{X}}$ is then obtained as $\tilde{\mathbf{X}} = \mathbf{X}\mathbf{V}\Lambda^{-1/2}$. Now $\tilde{\mathbf{X}}^T\tilde{\mathbf{X}} = \Lambda^{-\mathbf{1/2}}\mathbf{V}^T\mathbf{X}^T\mathbf{X}\mathbf{V}\Lambda^{-\mathbf{1/2}} = \mathbf{I}$. $\tilde{\mathbf{X}}$ can also be directly obtained from a SVD of $\mathbf{X} = \mathbf{U}\mathbf{S}\mathbf{V}^T$, where $\Lambda = \mathbf{S}^2$ and $\tilde{\mathbf{X}} = \mathbf{U}$. The dimensionality of the data can be reduced in the same step by retaining only the first D eigenvectors of $\mathbf{U}$.

$s = 1, \ldots, S$), the CCA solution simplifies considerably:

$$\begin{bmatrix} 0 & \mathbf{C_{12}} & \mathbf{C_{13}} & \ldots & \mathbf{C_{1S}} \\ \mathbf{C_{21}} & 0 & \mathbf{C_{23}} & \ldots & \mathbf{C_{2S}} \\ \vdots & & \ddots & & \vdots \\ \mathbf{C_{S1}} & \ldots & \ldots & \ldots & 0 \end{bmatrix} \mathbf{A} = \Lambda \mathbf{A}, \tag{3.3}$$

where $\mathbf{A}$ stands for the row-concatenated projection matrices $\mathbf{A_s}$.

We now add $\mathbf{IA}$ to both sides of Eq. (3.3) because this allows us to express the cross-covariance matrix as the covariance matrix of the concatenated, whitened datasets:

$$\begin{bmatrix} \mathbf{I} & \mathbf{C_{12}} & \mathbf{C_{13}} & \ldots & \mathbf{C_{1S}} \\ \mathbf{C_{21}} & \mathbf{I} & \mathbf{C_{23}} & \ldots & \mathbf{C_{2S}} \\ \vdots & & \ddots & & \vdots \\ \mathbf{C_{S1}} & \ldots & \ldots & \ldots & \mathbf{I} \end{bmatrix} \mathbf{A} = \begin{bmatrix} \mathbf{X_1}^T \\ \mathbf{X_2}^T \\ \vdots \\ \mathbf{X_S}^T \end{bmatrix} \begin{bmatrix} \mathbf{X_1} & \mathbf{X_2} & \ldots & \mathbf{X_S} \end{bmatrix} \mathbf{A} = (\Lambda + \mathbf{I})\mathbf{A} \tag{3.4}$$

The addition of the identity matrix leaves the eigenvectors unchanged and increases all eigenvalues by one. Eq. (3.4) describes a conventional PCA, completing the derivation.

In the journal article presented in this chapter, we solved $\mathbf{X'X'}^T = \mathbf{U}\Lambda\mathbf{U}^T$, while Eq. (3.4) solves $\mathbf{X}^T\mathbf{X} = \mathbf{A}(\Lambda + \mathbf{I})\mathbf{A}^T$ (note the location of the transposes). $\mathbf{A}$ is therefore similar to what we previously called “weights", and we obtain the eigenconnectivities by projecting and normalizing: $\mathbf{U} = \mathbf{XA}(\Lambda + \mathbf{I})^{-1/2}$. Or, simpler, by an SVD of the whitened data: $\mathbf{X} = \mathbf{U}\sqrt{(\Lambda + \mathbf{I})}\mathbf{A}^T$. The eigenconnectivities obtained from concatenated subject principal components (instead of dFC estimates) thus correspond to a subspace common to all subjects as obtained by gCCA.

3.2.2 gCCA eigenconnectivities

To compare PCA with gCCA, we applied a PCA to each subject's dFC matrix and (arbitrarily) retained the first 19 ± 2 or 5 ± 1 eigenvectors, corresponding to 90% and 50% of explained variance, respectively (see Fig. 3.9b for the eigenvalue spectrum of the dFC matrix of an example subject). In Figs. 3.9c and e we show the first 10 eigenconnectivites obtained from whitened data. The eigenconnectivities strongly resemble those obtained without whitening (Figs. 3.9d and f), suggesting that the PCA patterns are not driven by subject-specific noise or dFC spikes, which would either not be retained after the subject-wise PCA or are unlikely to be shared among subjects. The eigenconnectivities are well preserved with as little as 50% of retained variance from the subject-level PCAs. With 90% of retained variance the eigenconnectivites are still largely similar, but less so than with 50% of retained variance, possibly because the whitening normalizes each subject's data to unit variance, which means that the 19^{th} principal component has a weight equal to the first one.

3.3 Outlook

In this chapter we presented an approach for extracting co-varying functional connections in multi-subject dFC datasets based on PCA. PCA is a classical linear analysis technique and is easy to interpret (because of its linearity), fast (because it is based on an eigenvalue decomposition), easy to implement

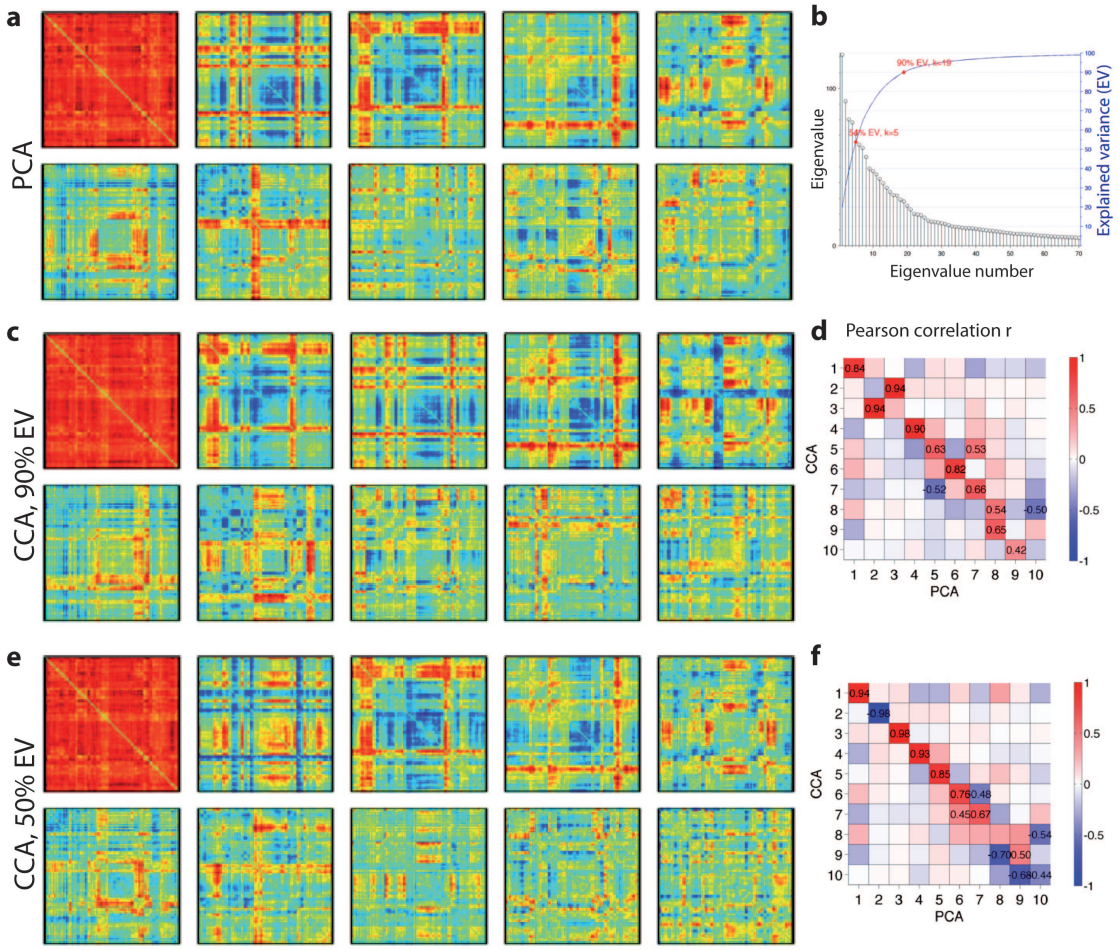


Figure 3.9 – (a) Eigenconnectivities obtained from concatenated dFC matrices, as also shown in Fig. 2.3, (b) Eigenvalues 1 to 70 out of 205 for a subject-wise PCA of dFC, (c) eigenconnectivities obtained from a multi-subject PCA with 90% of retained variance for each subject-wise PCA, (d) Pearson correlation r between the patterns shown in (a) and (c), the values of correlations $r > 0.4$ are indicated, (e) eigenconnectivities obtained from a multi-subject PCA with 50% of retained variance for each subject-wise PCA, (f) Pearson correlation between the patterns shown in (a) and (e).

and many extensions are built on it. The FC patterns we identified showed marked differences to those obtained from traditional, “average" FC analyses, and were robust to a two-level analysis, which first employed a whitening step. Our results also suggest altered FC dynamics in RRMS, but will require further validation in independent data sets.

Similar to our idea, Bassett *et al.* (2014) recently proposed to uncover sets of connections with similar dynamics by temporally correlating dFC of single-subjects. Both approaches thus aim to identify groups of co-varying connections. For us, the temporal covariance of dFC builds the basis for this, and for Bassett *et al.* (2014) it is the temporal correlation of dFC. Most importantly, while we apply PCA to uncover the groups, they directly study the thresholded dFC correlation matrix of size $(N^2 - N)/2 \times (N^2 - N)/2$.

Structured covariance of dynamic networks has recently also been revealed on much faster time

scales in human EEG data, supporting a neuronal basis for FC dynamics (Mehrkanoon *et al.*, 2014; Spencer *et al.*, 2013). Both Mehrkanoon *et al.* (2014) and Spencer *et al.* (2013) derived a low-dimensional orthogonal space of resting-state network dynamics by applying PCA to networks estimated from the phase synchronization of neural activity. Mehrkanoon *et al.* (2014) identified several networks with distinct topologies and frequency spectra and Spencer *et al.* (2013) accurately classified a motor task from the temporal expression of the components.

The neurological significance of the individual components is more difficult to grasp because of the orthogonality constraint. The eigenconnectivities span a low-dimensional space in which dFC unfolds, yielding an efficient representation of network dynamics, but orthogonality is not a natural assumption for the basic modes of brain function. In this work we also didn't explore the suitability of PCA *per se* for describing FC dynamics. Both of these points will be addressed in the next chapter.

4 Disentangling the temporal composition of dynamic brain networks

Is dFC well characterized by a mixture of FC patterns at every window? Or is it better characterized by alternating states? In the previous chapter we proposed to identify recurring network configurations using PCA, which decomposes dFC into the superposition of multiple connectivity patterns (Fig. 4.1). Alternatively, Allen *et al.* (2014) proposed to cluster dFC, which then describes dFC as a succession of separated states. Because the choice of decomposition impacts both the conceptual understanding of dFC and the interpretation of the estimated patterns, the goal of this chapter is to explore which temporal characterization is more appropriate.

To this aim, we apply matrix factorizations subject to variable constraints in the temporal domain to simulated, task-based and resting-state dFC of young, healthy subjects. Specifically, we decompose a multi-subject dFC matrix **C** into connectivity patterns **D** and associated temporal weights **A**: $\mathbf{C} \approx \mathbf{DA}$. K-means clustering allows only a single non-zero, positive weight at each time point (i.e., only one active FC pattern), while PCA allows the weights of all FC patterns to be non-zero at the same time (i.e., combined expression of FC patterns). Our results suggest that dFC during subject-driven tasks is well described by alternating FC states and that dFC during rest is better described by multiple FC patterns that overlap at each time instance.

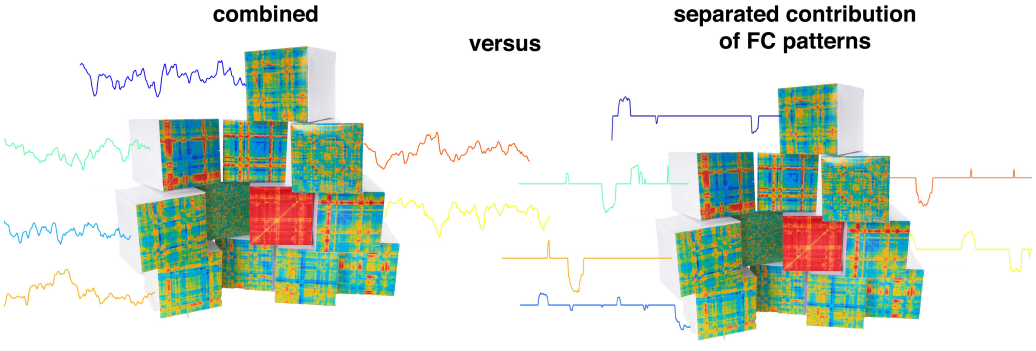


Figure 4.1 – dFC is decomposed into building blocks and associated temporal weights. Depending on the decomposition, dFC is modelled as a combination of all FC patterns or as the expression of a single FC pattern. Photograph courtesy of M. Leonardi.

4.1 Journal article

Submitted to: Human Brain Mapping (under revision)

Disentangling dynamic networks by sparse combinations of functional connectivity patterns

Nora Leonardi,[1,2] William R. Shirer,[3] Michael D. Greicius,[3] and Dimitri Van De Ville[1,2]
[1] Institute of Bioengineering, École Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland; [2] Department of Radiology and Medical Informatics, University of Geneva, Geneva, Switzerland; [3] Department of Neurology and Neurological Sciences, Stanford University, Stanford, USA

Resting-state functional connectivity (FC) is highly variable across the duration of a scan. Groups of co-evolving connections, or reproducible patterns of dynamic FC (dFC), have been revealed in fluctuating FC by applying unsupervised learning techniques. Based on results from k-means clustering and sliding-window correlations, it has recently been hypothesized that dFC may cycle through several discrete FC states. Alternatively, it has been proposed to represent dFC as a linear combination of multiple FC patterns using principal component analysis. As it is unclear whether sparse or non-sparse combinations of FC patterns are most appropriate, and as this affects their interpretation and use as markers of cognitive processing, the goal of our study was to evaluate the impact of sparsity by performing an empirical evaluation of simulated, task-based and resting-state dFC. To this aim, we applied matrix factorizations subject to variable constraints in the temporal domain and studied both the reproducibility of ensuing representations of dFC and the distribution of time-dependent weights. During subject-driven tasks, dFC was well described by alternating FC states in accordance with the nature of the data. The estimated FC patterns showed a rich structure with combinations of known functional networks enabling accurate identification of the three different tasks. A fourth FC state was apparent during one of the three tasks, suggesting subjects did not perform the task continuously. During rest, dFC was better described by multiple FC patterns that overlap. The executive control networks, which are critical for working memory, appeared grouped alternately with externally- or internally-oriented networks. These results suggest that combinations of FC patterns provide a meaningful way to disentangle resting-state dFC.

4.1.1 Introduction

Brain networks reconfigure over years (Dosenbach *et al.*, 2010), days (Bassett *et al.*, 2011) and seconds (Chang & Glover, 2010; Eryilmaz *et al.*, 2011; Cribben *et al.*, 2012; Ekman *et al.*, 2012). Fast, dynamic reconfigurations may occur in response to a changing external environment or spontaneously while a subject is at rest. In particular, spontaneous fluctuations in the temporal correlation of the BOLD activity of distinct brain regions have been observed in the default mode network (DMN) (Chang & Glover, 2010) and since been shown to also occur in many other large-scale networks (Smith *et al.*, 2012; Hutchison *et al.*, 2013b; Liu & Duyn, 2013; Allen *et al.*, 2014). Using unsupervised learning techniques, recent work has identified reproducible groups of functional connections that evolve in a similar manner, which we here call "functional connectivity (FC) patterns" (Eavani *et al.*, 2013;

Leonardi & Van De Ville, 2013b; Li *et al.*, 2014; Allen *et al.*, 2014). The occurrence of various FC patterns across time challenges the assumption of unique and stable correlations during rest (Chang & Glover, 2010; Allen *et al.*, 2014), and links between FC variability and neural activity have been reported (Chang *et al.*, 2013b; Allen *et al.*, 2013). Early work also suggests that these patterns are altered in several neuropsychological diseases (Jones *et al.*, 2012; Leonardi & Van De Ville, 2013b; Li *et al.*, 2014), highlighting the potential importance of dFC to provide insights that are complementary to traditional, static FC (see e.g. Hutchison *et al.*, 2013a, for a recent review).

Sliding-window correlation analysis has been the most commonly used approach to study dynamic FC (dFC) across time and Allen *et al.* (2014) combined it with k-means clustering to separate dFC into several so-called "FC states". In related work, Li *et al.* (2014) also used k-means clustering as a first step to identify FC patterns. Preliminary results also indicate that the application of k-means clustering to sliding-window correlations can successfully distinguish different tasks in individual subjects for single subjects (Gonzalez-Castillo *et al.*, 2013). Thus, dFC is sensitive to changes in cognitive states and provides enough information to reveal them. Based on these results it has been suggested that dFC may cycle through multiple discrete states during rest (Hutchison *et al.*, 2013a). We posit that clustering inevitably models dFC as a succession of states because its aim is to find a simple representation by approximating each windowed FC estimate by a single FC pattern. The simplicity of clustering makes it an attractive, yet also restrictive approach, and to our knowledge there is no study evaluating the importance of the implicit sparsity assumption.

Alternatively, it has been proposed to represent dFC as a linear combination of multiple FC patterns by applying principal components analysis (PCA) (Leonardi *et al.*, 2013) and a tensorial extension for task data enabled accurate classification of two cognitive states (Leonardi & Van De Ville, 2013b). PCA is a more flexible approach, but it can be prone to overfitting. Again, we posit that PCA inevitably models dFC as the combination of all (orthogonal) patterns. To understand whether sparsity is relevant from a conceptual point of view and how the FC patterns should be interpreted, we performed an empirical study.

We evaluate the impact of sparsity by casting both k-means clustering and PCA[1] as matrix factorizations, which decompose dFC into components and associated time-dependent weights: In clustering, each dFC estimate is approximated by a single component (the cluster centroid) with a weight of one, while in PCA each dFC estimate is approximated by a linear combination of all orthogonal components (Fig. 4.2). Thus, clustering assumes a maximally sparse model, while PCA assumes no sparsity (but orthogonality). Matrix factorizations subject to variable sparsity constraints can be seen as generalizations of both k-means clustering and PCA, and have previously been shown to well describe BOLD activity and static FC (Lee *et al.*, 2011, 2013; Eavani *et al.*, 2012). Here, by comparing sparse and non-sparse linear combinations of FC patterns we explore whether dFC is better described by a succession of separated FC states or by the joint expression of multiple FC patterns. We first simulated data to test the feasibility of the proposed approach. Then, we decomposed dFC during three subject-driven cognitive states, which provide a more naturalistic view of continuous, cognitive processing than traditional task data. Importantly, because the cognitive state a subject is in was known, we could compare the estimated FC patterns and associated weights with the known experimental paradigm. Finally, we decomposed dFC during resting state.

[1] Technically, PCA decomposes a matrix only into (principal) components and it is the singular value decomposition (SVD) that decomposes it into components and weights. For centered data, the principal components of PCA are, however, equal to the left singular vectors of SVD. The weights can then be obtained by a projection of the data onto the components.

4.1.2 Methods

Participants and data acquisition

24 healthy, right-handed subjects (age range 18 - 30 years, 15 females) participated in this study. The data have previously been used to decode subject-driven cognitive states (Shirer *et al.*, 2012). The study was approved by the institutional review board of Stanford University and informed consent was obtained from each subject. Functional images were acquired on a 3.0T GE scanner (TR = 2s, TE = 30 ms, flip angle = 80°, 1 interleave, matrix size 64 x 64, FOV = 22 cm). A high-resolution structural scan was acquired using an axial 3D fast spoiled gradient recalled echo sequence (FSPGR; 162 slices, 0.86-mm^2 in-plane and 1-mm through-plane resolution, TR = 5.9 ms, TE = 2 ms, flip angle = 15°, FOV = 22 cm). Participants were scanned for 10 minutes each while resting and completing three subject-driven tasks: counting backwards from 5000 in steps of 3 (subtraction task), recalling the events of their day (episodic memory task), and silently singing a song (music task). No stimulus was presented and the investigator only marked the start and end of each 10-min scan. Subjects were instructed to keep their eyes closed. A debriefing confirmed that all subjects stayed awake and were able to perform the tasks throughout the scan.

Preprocessing

Data were preprocessed and analyzed using FMRIBs Software Library (FSL version 4.1). The first six volumes were discarded to allow the MR signal to equilibrate. Data were corrected for motion, normalized to the MNI template, and smoothed with a 6-mm Gaussian kernel. Subject's heart rate and respiration rate were monitored during the scan and regressed out from the fMRI data, together with motion parameters, cerebrospinal fluid and white matter signals, and a brain-averaged signal. Finally, data were high-pass filtered (>0.01 Hz). BOLD activity was averaged within 78 functional ROIs, which were previously defined from thresholded independent components estimated from 14 of the 24 subjects and one other subject (`http://findlab.stanford.edu/functional_ROIs.html`, cerebellar ROIs were excluded; Shirer *et al.*, 2012).

Dynamic functional connectivity analysis

dFC estimation

We estimated dFC between all N=78 ROIs using sliding-window correlations (Chang and Glover, 2010). We used a window size of 30 TRs (60 s) and shifted the window by 5 TRs (10 s) for each estimation, resulting in $W = 53$ windows for each subject. These choices are similar to values reported in the literature (Handwerker *et al.*, 2012; Hutchison *et al.*, 2013b; Leonardi & Van De Ville, 2013b; Allen *et al.*, 2014) and Shirer *et al.* (2012) reported that different cognitive states were well distinguished with 60s of data.

We concatenated the vectorized, upper triangular parts of all 53 correlation matrices to build a connections × windows matrix $\mathbf{C_s} \in \mathbb{R}^{E \times W}$, where $E = N(N-1)/2$ (Fig. 4.2a). Finally, we subtracted the mean from each row of $\mathbf{C_s}$ and concatenated the centered $\mathbf{C_s}$ across subjects to form the connections × (windows × subjects) matrix $\mathbf{C} \in \mathbb{R}^{E \times T}$. We centered each connection to exclude any bias from inter-subject variability in average FC (Leonardi & Van De Ville, 2013b), and this step improved the correct identification of the three cognitive states in task-based dFC (adjusted rand index without centering = 0.27 and with centering = 0.74, where a value of 0 corresponds to a random labelling and a

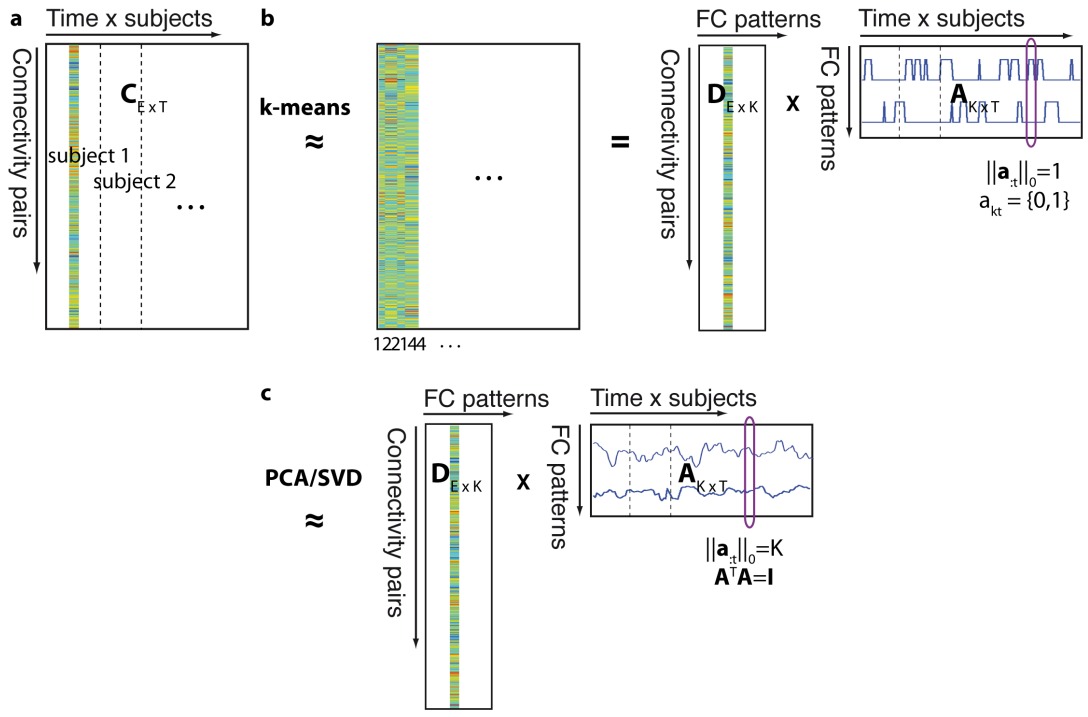


Figure 4.2 – (a) dFC is temporally concatenated across multiple subjects to form matrix **C**. (b) k-means clustering separates the data into K clusters, thereby approximating **C** by a succession of the cluster centroids, e.g. the first dFC estimate is approximated by centroid 1, the second and third by centroid 2, ... This can be seen as decomposing the dFC matrix **C** into K FC patterns (**D**) and associated sparse weights (**A**), where only one FC pattern may have a non-zero weight at each instance in time ($||\mathbf{a}_{:t}||_0 = 1$). (c) PCA or SVD decomposes the dFC matrix into K FC patterns and associated non-sparse, orthogonal weights ($||\mathbf{a}_{:t}||_0 = K$).

value of 1 to perfect labelling).

DFC matrix factorization

To compare k-means clustering and PCA/SVD we formulate both as matrix factorizations: $\mathbf{C} \approx \mathbf{DA}$. That is, the dFC matrix **C** is represented as the multiplication of K FC patterns, stored as the columns of the matrix $\mathbf{D} = [\mathbf{d}_{:1}, \mathbf{d}_{:2}, \ldots \mathbf{d}_{:K}] \in \mathbb{R}^{E \times K}$, and their associated time-dependent weights (i.e. their importance in representing dFC at each window) in $\mathbf{A} = [\mathbf{a}_{:1}, \mathbf{a}_{:2}, \ldots \mathbf{a}_{:T}] \in \mathbb{R}^{K \times T}$, where the index $_{:1}$ refers to the first column vector of a matrix. The problem of jointly learning **D** and **A** can be formulated as

$$\underset{\mathbf{D},\mathbf{A}}{\operatorname{argmin}} ||\mathbf{C} - \mathbf{DA}||_F^2 \tag{4.1}$$

where argmin indicates that **D** and **A** are chosen such as to minimize the approximation error $||\mathbf{C} - \mathbf{DA}||_F^2$, and $||\mathbf{X}||_F^2 = \sum_{ij} x_{ij}^2$ is the Frobenius norm, which is defined as the sum of all entries squared (matrix ℓ_2-norm). K-means clustering and truncated SVD represent two special cases of the more general Eq. (4.1).

K-means clustering finds exclusive and exhaustive cluster assignments $\mathscr{M}$ that minimize the within-cluster distances. For example, for a squared Euclidean distance, it solves $\sum_{k=1}^{K} \sum_{\mathbf{c}_t \in \mathscr{M}_k} ||\mathbf{c}_t - \mathbf{m}_k||^2$, where $\mathbf{c}_t \in \mathscr{M}_k$ indexes all networks assigned to cluster k, and $\mathbf{m}_k$ is the associated cluster center

that equals the average of all networks assigned to that cluster. We can write $\sum_{k=1}^{K}\sum_{\mathbf{c}_t\in\mathcal{M}_k}||\mathbf{c}_t-\mathbf{m}_k||^2$ as a matrix factorization analogue to Eq. (4.1), but subject to constraints on $\mathbf{A}$ (Fig. 4.2b):

$$\underset{\mathbf{D},\mathbf{A}}{\operatorname{argmin}}||\mathbf{C}-\mathbf{D}\mathbf{A}||_F^2,\quad s.t.\quad ||\mathbf{a}_{:t}||_0=1, a_{kt}=\{0,1\} \tag{4.2}$$

The constraint $||\mathbf{a}_{:t}||_0=1$ implies that each dFC network is approximated by a single FC pattern $\mathbf{d}_{:k}$ (the ℓ_0–norm counts the number of nonzero entries), and $a_{kt}=\{0,1\}$ that the associated weights are binary. Other distance functions, such as one minus the correlation or one minus the cosine, are less restrictive and allow a_{kt} to vary positively. We use k-means clustering with one minus the correlation distance as this is insensitive to offsets and changes in scaling, but sensitive to changes in topology.

Truncated SVD corresponds to other constraints on $\mathbf{A}$ (Fig. 4.2c):

$$\underset{\mathbf{D},\mathbf{A}}{\operatorname{argmin}}||\mathbf{C}-\mathbf{D}\mathbf{A}||_F^2,\quad s.t.\quad ||\mathbf{a}_{:t}||_0=K, \mathbf{A}^T\mathbf{A}=\mathbf{I}, \tag{4.3}$$

where $\mathbf{I}$ is the identity matrix, i.e., $\mathbf{A}$ is orthogonal, and $\mathbf{D}=\mathbf{C}\mathbf{A}^T$ are the first few eigenvectors of the connection-wise covariance matrix. $||\mathbf{a}_t||_0=K$ indicates that each dFC network is approximated as a combination of all FC patterns. This constraint does not need to be imposed, but we include it for the sake of comparison between the two approaches.

K-means clustering and truncated SVD can be generalized by k-SVD (Aharon *et al.*, 2006):

$$\underset{\mathbf{D},\mathbf{A}}{\operatorname{argmin}}||\mathbf{C}-\mathbf{D}\mathbf{A}||_F^2,\quad s.t.\quad ||\mathbf{a}_{:t}||_0\leq S, \tag{4.4}$$

where the case $S=1$ generalizes k-means clustering (without constraining the values to be binary or positive) and the case $S=K$ generalizes PCA (without enforcing orthogonality, which, however, still leads to a solution that spans the same subspace: i.e., one can be represented as a linear combination of the other). Using this formulation only the sparsity of the weights separates the two approaches. We used a fast implementation of the optimization toolbox SPAMS for solving Eq. (4.4) (SPArse Modeling Software, `http://spams-devel.gforge.inria.fr/`; Mairal *et al.*, 2010).

Simulated and null data

We generated data under a spatiotemporal separability assumption by multiplying $K=3$ FC patterns with random time courses that corresponded either to the case of separated or joint expression. Specifically, for each subject $s=1,2,\ldots,24$, we multiplied subject-specific FC patterns $\mathbf{D}_s$ with randomly generated time courses $\mathbf{A}_s$: $\mathbf{C_s}=\mathbf{D_s}\mathbf{A_s}$, where $\mathbf{C_s}$ is the connection-by-window subject-specific simulated data, $\mathbf{D_s}$ contained the $K=3$ FC patterns on its columns, and $a_{kt}\sim|\mathcal{N}(0,1)|$. The subject-specific FC patterns $\mathbf{D_s}$ were estimated from average FC during the three subject-driven cognitive states subtraction, memory and music (Supplementary Fig. B.1). This incorporated a large amount of inter-subject variability as can be expected in real fMRI data. To simulate the expression of separated patterns, we randomly set two entries of $\mathbf{A_s}$ per column to zero ($S=1$) (Supplementary Fig. B.1). Finally, we added zero-mean Gaussian noise with $\sigma=0.02$ to each subject's data matrix $\mathbf{C_s}$.

To model the case of static FC across time, we multiplied single, subject-specific FC patterns with random time courses and added noise. Fluctuations in FC in this data set would thus merely reflect varying strengths of the same underlying network topology.

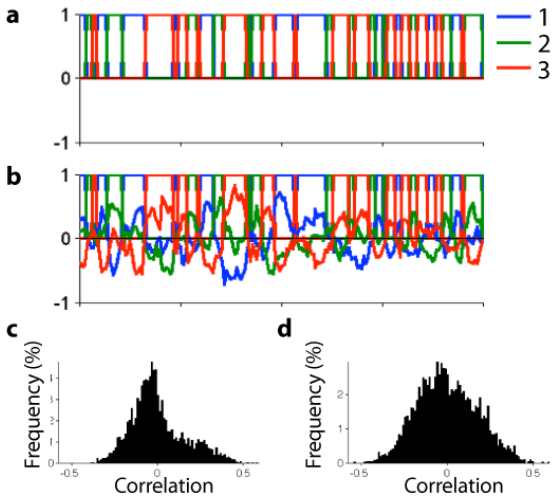


Figure 4.3 – (a) K-means clustering assigns each dFC network to one of three clusters. (b) This assignment is guided by the similarity of each dFC network to all three FC patterns. We interpret the similarity of each dFC network to all FC patterns as a non-sparse weight matrix $\mathbf{A}^*$. (c) The histogram of $\mathbf{A}^*$ distinguishes separated (left) and joint (right) expression of FC patterns.

Evaluation of matrix factorizations

Comparison of algorithms. We compared k-means clustering, truncated SVD and k-SVD (with $S = 1$ and $S = K$) by assessing how well they recovered the true, underlying FC patterns used for the simulations with K set to the true number of 3 (we study how to estimate K from the data below). Because $\mathbf{D_s}$ differed slightly between subjects, we used $\mathbf{D_s}$ averaged across all subjects as the true FC patterns we would like to recover. We quantified recovery by estimating the correlation between the true and the estimated FC patterns. For k-SVD, we retained the absolute values of the correlation coefficients because the sign of the FC patterns is arbitrary. For PCA, we estimated only $K - 1 = 2$ FC patterns because the components represent axes in a lower-dimensional space and two components can describe three, correlated FC patterns. The recovery of the third FC pattern was assessed by looking at negative correlations (i.e., one estimated FC pattern recovered two true FC patterns, one of which with a flipped sign). Given the cross-correlation matrices, we matched the true FC patterns with the estimated ones using the Hungarian algorithm (Kuhn, 1955; Munkres, 1957). The Hungarian algorithm finds the matching that maximizes the sum of similarity measures (here correlation coefficients) between all corresponding pairs. We report the average correlation coefficient of the matched pairs across 100 simulations. Because k-means clustering performed best overall in recovering the true FC patterns we retained only k-means clustering for the following steps.

Number of FC patterns K. We chose the number of FC patterns (or clusters) K by evaluating the reproducibility of the FC patterns in split-half resamplings. The subjects' dFC networks were randomly split into two, independent data sets of 12 subjects each. K-means clustering was applied separately to each data set, the cross-correlation between the FC patterns of the two data sets estimated, and the best matching estimated using the Hungarian algorithm. We define the reproducibility as the smallest correlation coefficient between all corresponding pairs (least reproducible FC pattern). We repeated all steps for 24 random splits.

Separated versus joint expression of FC patterns. K-means clustering assigns each dFC network to a single FC pattern by picking the one it is most similar to. That is, k-means clustering "sees" how

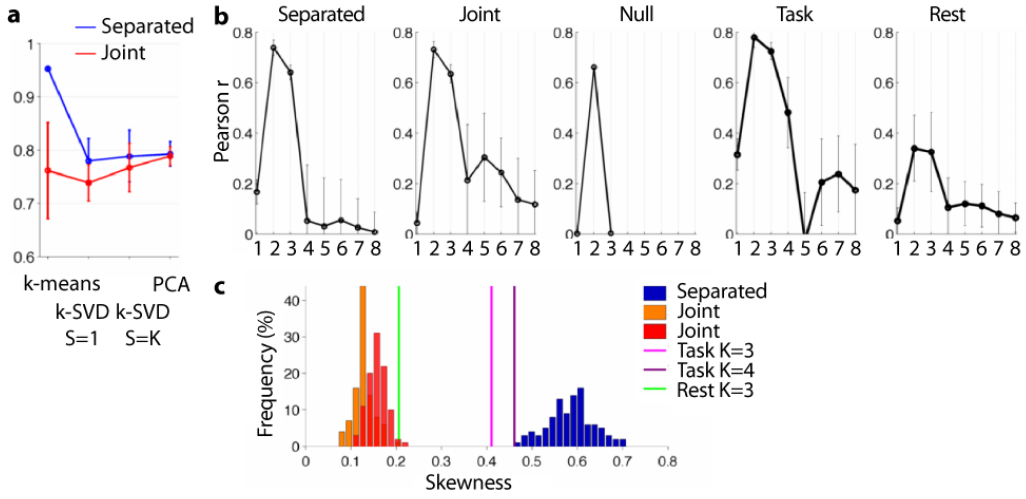


Figure 4.4 – (a) Average correlation coefficient of true FC patterns with those estimated using k-means clustering (K=3), sparse matrix factorizations (K=3), and PCA (K=2) for both separated and joint expression of FC patterns. Errorbars represent standard deviation across simulations. (b) Split-half reproducibility for $K = 1,2,\ldots,8$ for simulations (separated, joint and null), subject-driven task-based and resting-state dFC. Errorbars represent standard deviation across splits. (c) Skewness of time-dependent weights $\mathbf{A}^*$ for simulations and subject-driven task-based and resting-state dFC.

well each dFC network is approximated by all FC patterns, but then binarizes this information by retaining only the one with maximal similarity. We calculated the correlation of each dFC network to all K FC patterns to generate a non-sparse weight matrix $\mathbf{A}^*$, and assessed the asymmetry of the histogram of all correlation coefficients by calculating its skewness (Fig. 4.3). The skewness of normally distributed data is zero, and non-zero skewness indicates asymmetry.

4.1.3 Results

Simulated dFC

First, we compared how well k-means clustering, PCA and k-SVD (S=1 ≈ k-means clustering, and S=K ≈ PCA) recovered the true, underlying FC patterns in simulations. For the separated expression of FC patterns (favoring the sparsity of k-means), k-means clustering clearly outperformed all other algorithms (average correlation coefficient 0.95, versus < 0.8 for all other algorithms; Fig. 4.4a). For the joint expression (favoring the overlap of truncated SVD), all algorithms performed comparably (average correlation coefficients 0.74 – 0.79; Fig. 4.4a). Because (i) k-means clustering performed best for the separated expression and comparably for the joint expression, (ii) it avoids the ambiguity of the sign of the FC patterns (in k-SVD and truncated SVD the sign of the FC patterns is arbitrary), and (iii) it avoids the choice of the parameter S necessary for k-SVD, we estimated all further FC patterns with k-means clustering.

The split-half reproducibility for different numbers K of FC patterns showed a clear drop once K exceeded the number of underlying FC patterns for both the separated and joint expression of FC patterns (i.e., $K \geq 4$; Fig. 4.4b). In the null data, only $K = 2$ FC patterns were reproducible, corresponding to the trivial cases of above and below average FC.

The histogram of the weights $\mathbf{A}^*$ was highly asymmetric for the case of separated expression

(large, positive skewness; Fig. 4.3c and Fig. 4.4c). That is, each dFC network resembled only one FC pattern strongly at each time point and thus many correlations were close to zero. The histogram of the weights $\mathbf{A}^*$ was only weakly asymmetric for the simulations of joint expression (small, positive skewness; Fig. 4.3d and Fig. 4.4c). Because each dFC network resembled all FC patterns to some extent, correlations were more evenly distributed.

The split-half reproducibility of the FC patterns and the shape of the histogram of the time-dependent correlations with the FC patterns $\mathbf{A}^*$ could thus provide information on the underlying type of data generation.

Subject-driven task-based dFC

Fig. 4.5a shows average dFC for each subject-driven task: The subtraction task was characterized by strong FC in the visuospatial network, the episodic memory task by strong FC in the vDMN (retrosplenial cortex/medial temporal lobe network), and the music task by strong FC in the language network and the dDMN. These results are in line with those of Shirer *et al.* (2012).

Fig. 4.4b shows the split-half reproducibility for task-based dFC, which starts to drop at K=3, and is small after K=4. This suggests 3 or 4 FC patterns best describe the data. The asymmetry of the histogram of the weights $\mathbf{A}^*$ was close to that for simulated separated expression, correctly suggesting that a separated expression of the FC patterns well describes the data (Fig. 4.4c).

In Fig. 4.5b we illustrate the FC patterns estimated for K=3 and their associated time-dependent weights $\mathbf{A}^*$. The three patterns strongly ressembled the average dFC of each subject-driven task (Pearson correlation coefficients $r = 0.99, 0.98$, and 0.97) and the weights separated the three cognitive states well (adjusted rand index of k-means labelling = 0.74).

The fourth FC pattern showed two large groups of co-evolving connections: one centered on externally-oriented networks (salience, visuospatial (VS) and primary sensory networks) and one centered on internally-oriented networks (default mode network, DMN; Fig. 4.5c). The fourth FC pattern was specific to the music condition: it correlated with average FC during the music task and its weights indicated that it was predominantly – and consistently across subjects – expressed during this task.

Resting-state dFC

For resting-state dFC, split-half reproducibility dropped after $K = 3$ and the asymmetry of the histogram of the weights $\mathbf{A}^*$ overlapped with the simulated joint expression. This suggests that three FC patterns that are jointly expressed best describe resting-state dFC (Figs. 4.4b and c).

The estimated FC patterns showed distinct, large-scale network topologies. FC pattern 1 revealed two groups of co-evolving connections: a first group composed of the primary sensory (auditory, motor and visual), salience and VS (or dorsal attention network) networks, and a second group of bilateral executive control networks (ECN) and the DMN. FC pattern 2 showed a topology opposite to the one of FC pattern 1. We note again that each dFC time series was centered, which means that these FC patterns reflect excursions around average FC. Thus, FC pattern 2 does not indicate positive correlation between externally- and internally-oriented networks, but rather reduces the extent of their separation into distinct groups when added to average FC (and combined with the other FC patterns). FC pattern 3 highlighted a salience-VS-RECN group. A second group of connections in FC pattern 3 showed visual and language ROIs connected to the DMN.

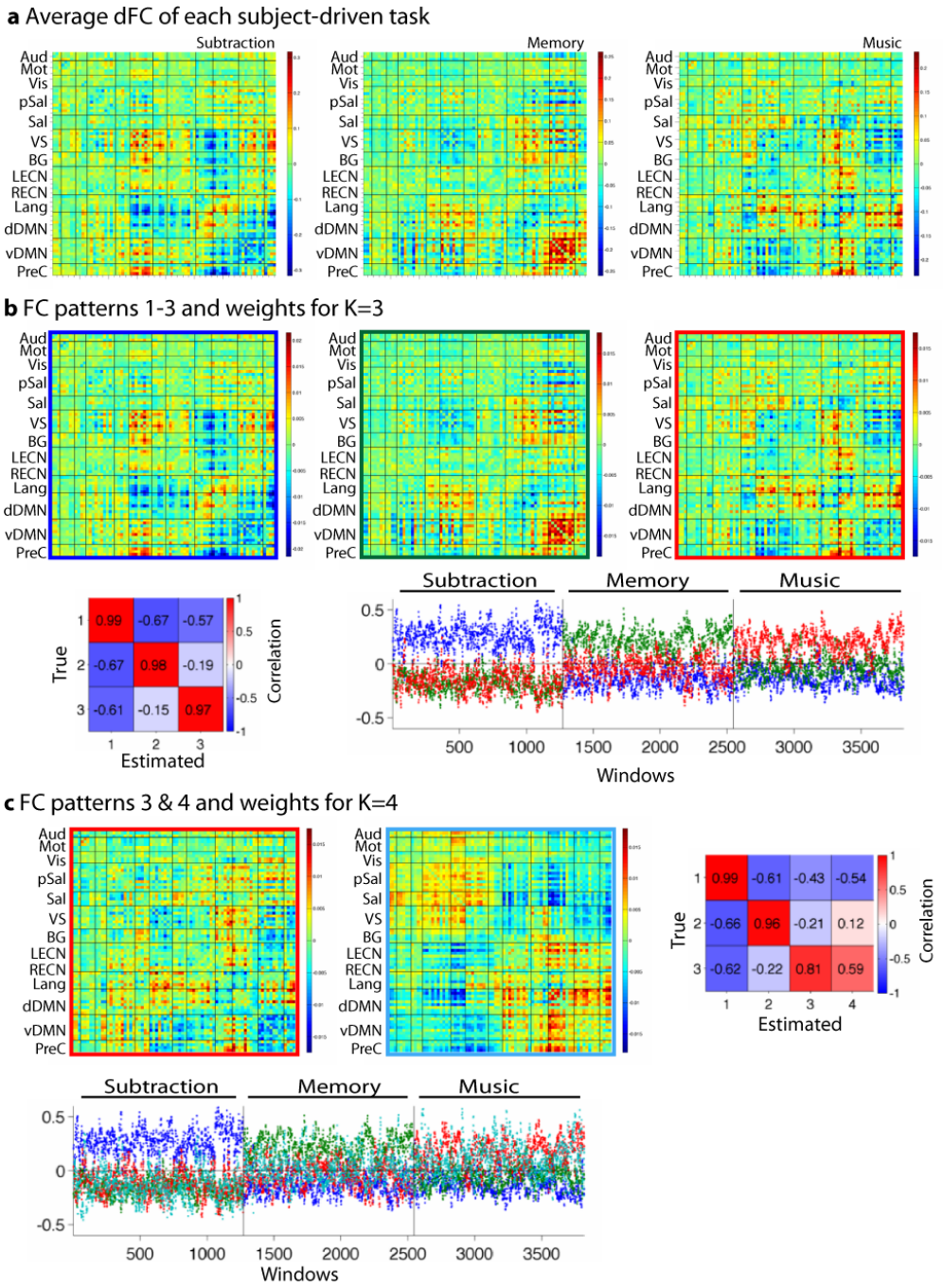


Figure 4.5 – (a) Average dFC of the three subject-driven tasks: subtraction, memory and music. (b) Estimated FC patterns (K=3), correlation coefficient with average FC of each task, and correlation coefficients of dFC networks with all 3 FC patterns across time and subjects (coloring as boxes around FC patterns). The x axis is arranged by task, rather than by subject. (c) Same as b for K=4 (only the two music-related FC patterns are shown as the other two do not change).

4.1.4 Discussion

We revisited k-means clustering and PCA, two recently proposed approaches to identify reproducible network configurations from windowed FC, to study the importance of temporal sparsity for the

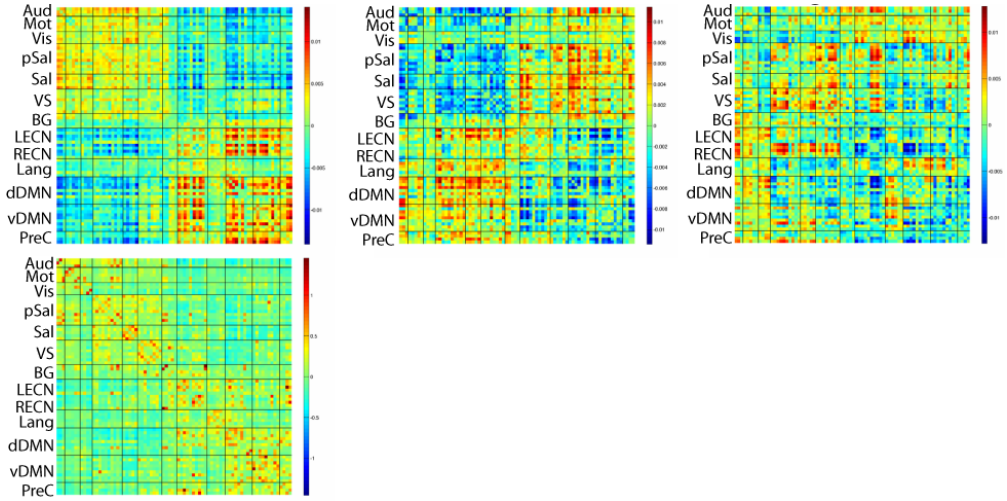


Figure 4.6 – FC patterns estimated during resting-state (top) and FC averaged across all windows and subjects (bottom).

representation and interpretation of dFC. To compare k-means clustering to PCA, we formulated it as a sparse matrix factorization. For data simulated according to both separated and joint expression of FC patterns, k-means clustering showed the best recovery of the underlying FC patterns overall. We then evaluated split-half reproducibility to identify the number of FC patterns from the data, and the asymmetry of the histogram of the time-dependent weights to assess whether the FC patterns occurred exclusively or jointly over time. Our results indicate that separated FC patterns are a good representation of subject-driven task-based dFC, and that a joint expression of FC patterns is better suited for resting-state dFC.

Subject-driven task-based FC patterns

We first analyzed dFC during three free-streaming, subject-driven cognitive states. Shirer *et al.* (2012) previously showed that these states could be distinguished with high accuracy using supervised learning (i.e., training a classifier on data with known cognitive states) and Gonzalez-Castillo *et al.* (2013) recently reported that four different tasks could be identified from dFC alone by clustering windowed FC in individual subjects. We extend these findings by showing that subject-driven cognitive states can be successfully identified from dFC of multiple subjects using unsupervised learning. The correct identification of the three states from multi-subject dFC was greatly improved by first centering each subject's dFC time series so that fluctuations reflected deviations around the overall mean (or average FC). We were further able to identify the model parameters, such as number of FC patterns and their temporal expression, in a data-driven way. This indicated that task-based dFC was well modelled using separated FC patterns. The fact that FC fluctuations can be related to changes in subject-driven cognitive states is highly encouraging.

A fourth FC pattern was also reproducible across independent split-half datasets and this FC pattern occurred preferentially during the music state. Interestingly, the music state was the most difficult state to classify in the original study by Shirer *et al.* (2012; 60% versus > 80% accuracy for the other states). The occurrence of this FC pattern might suggest that subjects did not continuously perform the assigned task. The FC pattern itself shows a certain resemblance to the first resting-state FC pattern,

with, however, stronger contributions of the DMN and salience network.

Resting-state FC patterns

We next analyzed dFC during unconstrained resting-state and based on our findings we suggest that a joint expression of multiple FC patterns is a better adapted and more flexible view of dFC than a succession of k-means centroids.

Based on split-half reproducibility, we estimated three FC patterns that had a rich structure and showed distinct network configurations. For example, a fragmentation between salience and VS connections on one hand and DMN connections on the other hand in FC pattern 1. FC pattern 2, which encodes the opposite contrast, indicates that the separation between these two groups is less pronounced at other instances in time (average FC was removed before clustering the data so "positive correlation" in these FC patterns is not the same as in a standard correlation matrix). In FC pattern 3 the DMN was connected to the visual and language networks, but not the RECN, which was here connected to the salience and VS networks, suggesting different modes of inter-network interaction. Both Allen *et al.* (2014) and Liu & Duyn (2013) also observed correlations and co-activations, respectively, of DMN and visual regions. The ECN is critical for working memory and appeared grouped both with networks that scan the environment (salience and VS in pattern 3) and networks that ruminate over recent events or episodic memories (DMN, pattern 1). This suggests that flexible interactions between the ECN and externally- and internally- oriented networks are critical for monitoring, maintaining and manipulating recent information. Previous work has also suggested interactions between the ECN and VS (e.g. LaBar *et al.*, 1999) and between the ECN and the DMN (e.g. Spreng *et al.*, 2010).

We here observed only three FC patterns that were reproducible across independent data sets. Different, but generally larger, values have been reported in the literature for related studies. Several factors – apart from the decomposition approach – are likely to play a role. First, the way K is chosen. We used reproducibility, while in Leonardi & Van De Ville (2013b) and Allen *et al.* (2014) the choice of K was guided by accuracy (or approximation error) instead. Second, the number of included subjects. For example, Allen *et al.* (2014) clustered dFC of over 400 subjects into 7 states in initial work, and into 5 states in a follow-up study presented at OHBM with 23 subjects (Allen *et al.*, 2013). Third, the parameters of the acquisition (such as the TR and scan duration), the choice of ROIs and the window length, might all play roles. All studies mentioned above used (different) ROIs to reduce the spatial dimensionality of the data. As is the case for studies of static FC, different atlases hamper an easy comparison of results across studies. We hope that future work will help to clarify some of these issues. Lastly, because the correlation structure is averaged across the full window duration, changes in network topology that occur on faster time scales are unlikely to be picked up. In exploratory analyses we shortened the window length to 45s but did not see an indication for a larger number of reproducible FC patterns. We did not decrease the window length further as the identification of the cognitive states for the subject-driven task data deteriorated for 30s-long windows (adjusted rand index 0.33 for 30s, 0.63 for 45s, and 0.74 for 60s). Alternative techniques that circumvent sliding-window correlations might be better able to reveal faster network reorganizations (e.g. Smith *et al.*, 2012; Eavani *et al.*, 2013; Liu & Duyn, 2013).

4.1.5 Conclusion

dFC is a promising new measure of brain activity that can be extracted from fMRI data and has the potential to provide new insights into brain function. In this work, we investigated whether dynamic functional network configurations were better described as a succession of separated FC states or rather as a combination of multiple FC patterns. We presented a systematic and data-driven analysis in terms of split-half reproducibility and distribution of time-dependent weights for simulated and experimental fMRI data from subject-driven cognitive states and resting-state. These results showed that a small number of superposed FC patterns provided the best decomposition of resting-state dFC data. The meaning of decomposing resting-state dFC remains to be further validated and is an exploratory technique, as also pointed out by Hutchison *et al.* (2013a). The fact that different cognitive states could be identified from dFC alone is encouraging but generative and computational models are needed to understand how dFC emerges and is best analyzed. One promising avenue to improve the interpretation of dFC is to combine approaches as the ones presented here with concurrent EEG or other physiological, behavioral or clinical measures (Allen *et al.*, 2013; Chang *et al.*, 2013b; Preti *et al.*, 2014).

4.2 Conclusions and outlook

In this chapter we studied how well sparse (i.e., segregated) and non-sparse (i.e., overlapping) decompositions characterize the temporal organization of dFC. Our results suggest that k-means clustering reliably identifies FC patterns, both if they are separately and jointly expressed over time. The fact that k-means clustering can identify distinct clusters does not imply that they are successively expressed in time. Rather, by studying the similarity of each windowed FC matrix to all FC patterns, we showed that they are likely jointly expressed during resting-state. Approximating dFC by a single FC pattern thus appears too restrictive.

In our simulations, k-SVD performed less well than k-means clustering even though the former corresponds to a more flexible model. One reason for the poorer performance of k-SVD might be that the positivity constraint on the weights $\mathbf{A}$ is dropped. Unfortunately, we are not aware of any algorithms that can impose constraints on both the ℓ_0-norm and non-negativity. Alternatively, one could replace the ℓ_0 norm with an ℓ_1 norm ($||\mathbf{a}_{:t}||_1 = \sum_k |a_{kt}|$), which is the convexified form of the ℓ_0 norm and can be combined with a positivity constraint: $\underset{\mathbf{D},\mathbf{A}}{\operatorname{argmin}} ||\mathbf{C} - \mathbf{DA}||_F^2, \quad s.t. \quad ||\mathbf{a}_{:t}||_1 \leq \lambda, a_{k,t} \geq 0$. In exploratory analyses (without extensive tuning of the parameter λ), this alternative decomposition did, however, also not perform better than k-means clustering.

PCA performed well in the simulations of jointly expressed FC patterns, but the interpretation of the estimated FC patterns is harder than for k-SVD or k-means because of the orthogonality constraint. Because of this constraint, the FC patterns are axes in a low-dimensional space and 2 FC patterns sufficed to model 3 underlying (correlated) FC patterns. In resting-state data, where we do not have a ground truth for either the FC patterns or their temporal expression, they can therefore be hard to interpret. In the next chapter we explore dFC in task data, where the experimental paradigm together with the time-dependent weights of each FC pattern can help guide our interpretation.

In the previous chapter we noted that the eigenvalue spectrum of the PCA decayed slowly and that a large number of FC patterns would be necessary to explain a majority of the variance in the data. Based on our results and those of Li *et al.* (2014), we suggested that it might not be necessary to retain a large portion of the variance to study dFC and to identify relevant differences between different populations. The subject-specific PCA introduced in Section 3.2 and the task-based dFC studied in this chapter provide some further insights into the tradeoff between accuracy and reproducibility. First, a relatively small number of components explained a majority of the variance of single-subject dFC, but many components were necessary for multi-subject dFC. Second, the out-of-sample approximation was poor for both task-based and resting-state dFC (> 90% approximation error with 3 FC patterns during subject-driven tasks and > 94% during rest), but, despite this large approximation error, the three subject-driven tasks were well discriminated. This suggests that a small number of FC patterns is common across subjects and carries relevant information, while a large number of FC patterns captures subject-specific variability and noise. One explanation for this might be the intrinsic variability of sliding-window correlation analysis (Smith *et al.*, 2012; Handwerker *et al.*, 2012).

When estimating dFC using sliding-window correlation we have to choose a fixed window length. The length of the window is typically chosen based on a trade-off between the accuracy of the correlation estimate (not too short) and the sensitivity to changes in the network (not too long), and lengths between 30 and 60 seconds are common. We further discuss the influence of the window length on the appearance of spurious and real variability of sliding-window correlations in Appendix C.

Because each dynamic network is estimated from a temporal segment as long as the window, network reorganizations that occur on a faster time scale than the window duration cannot be detected. In Chapter 6 we explore an alternative approach to track brain network dynamics that has a better temporal resolution.

A consequence of the proposed matrix decompositions is that the identified FC patterns have global extent. This favors a description of global or overall features over one of local features (subsets of connections). As an alternative, we propose to analyze brain activity using localized FC patterns via a graph wavelet transform in the next chapter.

5 From global to local network analysis: the graph wavelet transform

PCA AND SPARSE MATRIX DECOMPOSITIONS can uncover connectivity patterns that reveal information about co-varying connections in the human brain. However, these patterns do not respect the underlying topology of the network because the correlation matrices are vectorized, and they have a global support covering all brain regions. In this chapter we propose to address these challenges by introducing a graph wavelet transform for dynamic brain networks. We extend the graph wavelet transform proposed by Hammond *et al.* (2010), which is based on a graph Fourier transform (Chung, 1997). The graph Fourier transform addresses the first challenge in that it respects the underlying graph topology, but its basis functions continue to be global in support (Figs. **??**a-c and f). In analogy to classical Fourier analysis such a basis is powerful for a global analysis but transient or local features in the data and structure at multiple scales are not easily revealed. The graph wavelet transform addresses these challenges through the use of graph-adapted basis functions at multiple scales and locations (Figs. **??**d and g). Again in analogy to classical wavelet analysis, graph wavelets act like a mathematical microscope and allow an efficient multiscale analysis. Classical wavelet analysis has in the past been successfully exploited for spatial and temporal analysis of fMRI data (Van De Ville *et al.*, 2004; Bullmore *et al.*, 2004; Breakspear *et al.*, 2006; Achard *et al.*, 2006).

The graph wavelet transform by Hammond *et al.* (2010) is defined for a single network only. We propose to generalize the transform to dynamic networks by identifying network components, which capture relevant aspects of the changing network architecture. To this aim, we first represent the set of dynamic networks as a third-order array, or tensor (Kolda & Bader, 2009; Dunlavy *et al.*, 2011). This representation gives us access to a wealth of methods from multilinear algebra. Then, we apply a tensor decomposition that generalizes SVD/PCA to higher orders to identify patterns in the underlying variable graph topology. New networks are finally constructed from these patterns for which we design graph wavelets.

Graph wavelets are well positioned to reveal local characteristics of the network itself or of data defined on the network. In the previous chapters we focused on the network structure but in this chapter we aim to exploit the multiscale property of the graph wavelets to study the co-activation of brain areas. The key idea here is that if the graph edges are defined from FC, rather than e.g. physical distance, we can expect neighboring brain areas (with respect to the underlying graph) to be activated at the same time. Because of the multiscale decomposition, we can analyze co-activations both at a fine scale, which reveals local co-activations (microscope with large magnification), and at a coarse

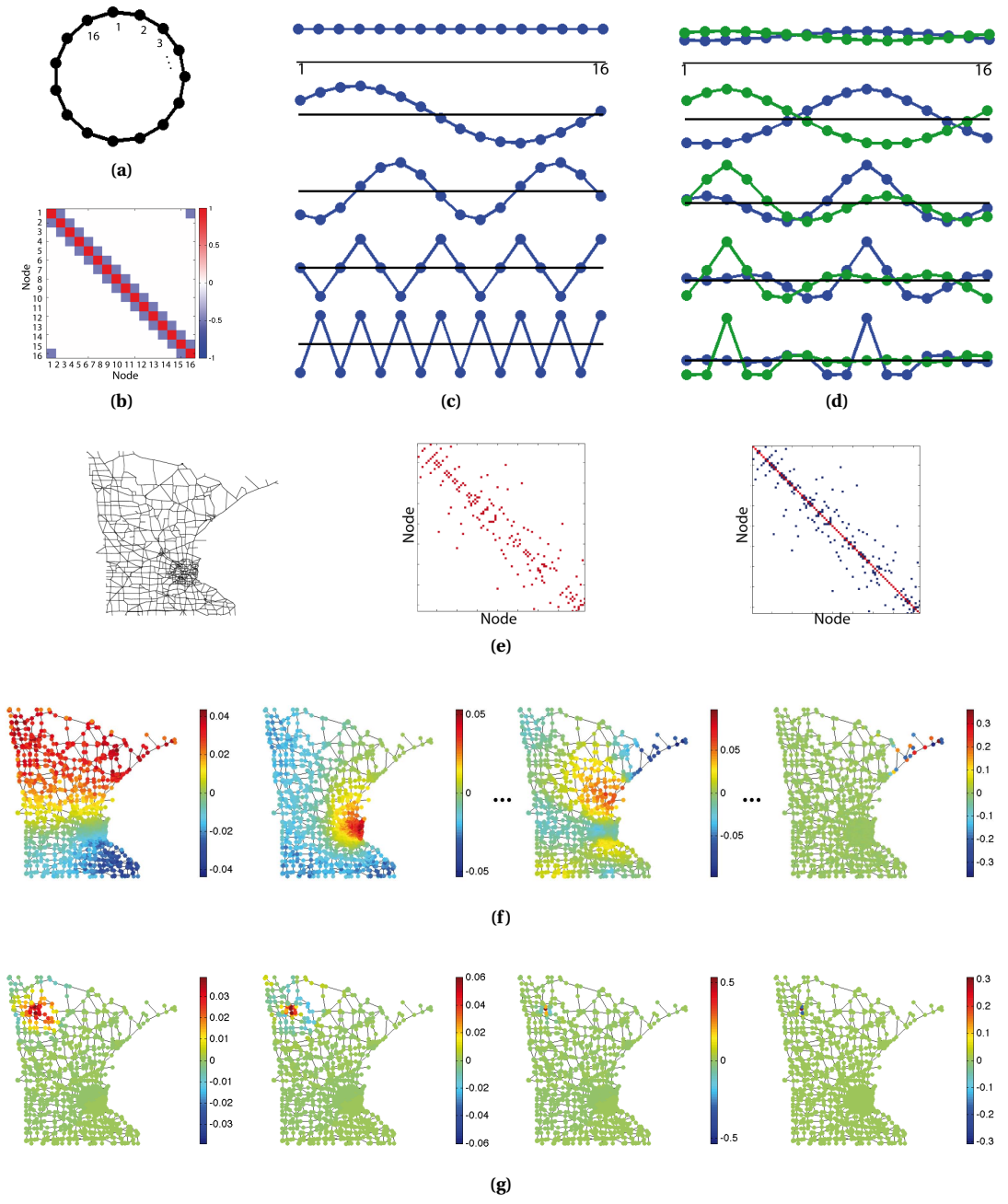


Figure 5.1 – (a) A cycle graph with 16 nodes, corresponding to a conventional, discrete grid with periodic boundary conditions. (b) The Laplacian matrix, a matrix representation of the graph. The diagonal contains (normalized) degrees of each network node and off-diagonal entries indicate connections between nodes. (c) Eigenvectors of the Laplacian matrix. These eigenvectors are the basis vectors of the discrete Fourier transform. (d) Two graph wavelets at multiple scales, centered at nodes 3 and 10. (e) Minnesota road graph and its matrix representations, the adjacency and Laplacian matrix (2640 nodes; Hammond *et al.*, 2010). (f) The eigenvectors of the Laplacian matrix are adapted to the irregular topology of the graph. (g) The graph wavelets dilate along the connections of the network. Left to right from coarse to fine scale, wavelet centered at an arbitrary node.

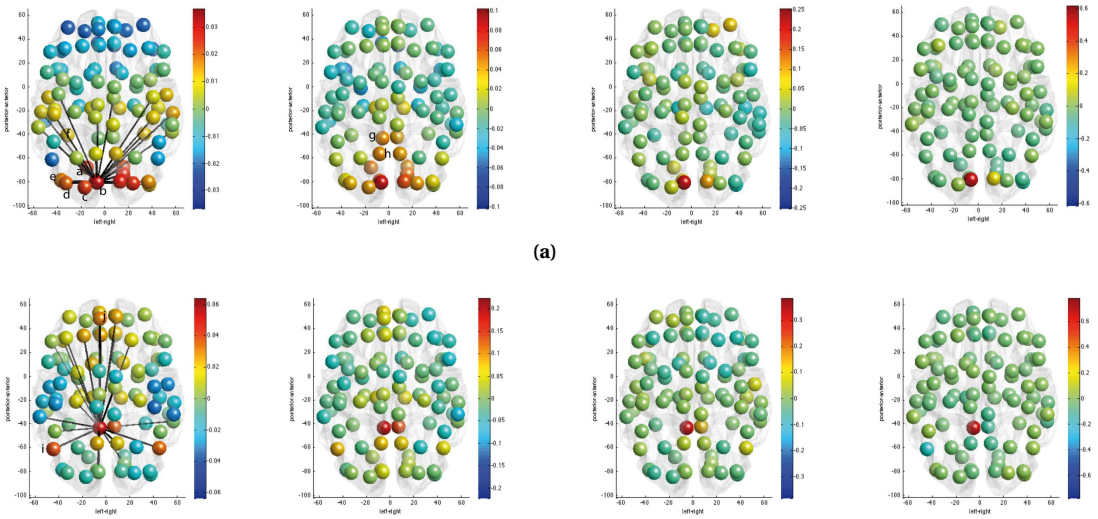


Figure 5.2 – Brain graph wavelets centered in the left cuneus (a) and left PCC (b) from coarse (left) to fine scales (right). Each sphere represents a brain region of the AAL atlas and colors indicate the value of the wavelet coefficient. The 25 strongest connections of the center node are shown at the coarsest scale. Note how the wavelets dilate to regions, which are part of the same "resting-state network", i.e., to regions of the visual cortex and default mode network, respectively. The labeled regions in the left hemisphere are: a, lingual gyrus; b, calcarine; c, superior occipital cortex (OC); d, middle OC; e, inferior OC; f, fusiform gyrus; g, PCC, h; precuneus; i, angular gyrus; j, medial prefrontal cortex.

scale, which reveals more global co-activations (low magnification). Fig. 5.2 shows how brain graph wavelets centered in the primary visual cortex and PCC, respectively, are well localized at fine scales, and dilate to functionally connected regions at coarse scales, that is to other visual regions for the wavelet centered in the cuneus; and to the angular gyri and medial prefrontal cortex for the wavelet centered in the PCC. Thus, we expect a basis that is adapted to the brain's functional wiring to be sensitive to the co-activation of functionally connected regions.

To explore this idea we analyze fMRI data acquired during alternating epochs of movies and rest, which robustly reorganizes average functional brain networks (Eryilmaz *et al.*, 2011; Richiardi *et al.*, 2011). We hypothesized that brain areas will co-activate at coarse scales during the movies condition when analyzed with graph wavelets defined for the movie network, but not for the resting state network (because the edges are not between brain areas that co-activate). In other words, we expect a signal that is smooth with respect to the movie network to not be smooth with respect to the resting-state network, and vice versa. By constructing graph wavelets adapted to either condition we hope to efficiently capture changes in the subject's cognitive states.

5.1 Journal article

Published in: IEEE Transactions on Signal Processing, 2013; 61(13):3357-3367.

Tight Wavelet Frames on Multislice Graphs

Nora Leonardi,[1,2] and Dimitri Van De Ville[1,2]

[1] Medical Image Processing Lab, Institute of Bioengineering, École Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland; [2] Medical Image Processing Lab, Department of Radiology and Medical Informatics, University of Geneva, Geneva, Switzerland

We present a framework for the design of wavelet transforms tailored to data defined on multislice graphs (i.e., multiplex or dynamic graphs). Graphs with multiple types of interactions are ubiquitous in real life, motivating the extension of wavelets to these complex domains. Our framework generalizes the recently proposed spectral graph wavelet transform (SGWT) (Hammond *et al.*, 2010), which is designed in the spectral (frequency) domain of an arbitrary finite weighted graph. We extend the SGWT to form a tight frame, which conserves energy in the wavelet domain, and define the relationship between conventional and spectral graph wavelets. We then propose a design for multislice graphs that is based on the higher-order singular value decomposition (HOSVD), a powerful tool from multilinear algebra. In particular, the multiple adjacency matrices are stacked to form a tensor and the HOSVD decomposition provides information about its third-order structure, analogous to that provided by matrix factorizations. We obtain a set of "eigennetworks" and from these graph wavelets, which exploit the variability across the graphs. We demonstrate the feasibility of our method (1) by capturing different orientations of a gray-scale image and (2) by decomposing brain signals from functional magnetic resonance imaging. We show its effectiveness to identify variability across graph edges and provide meaningful decompositions.

5.1.1 Introduction

Brain networks, human society or the internet are only some examples of systems that can be modelled as networks. Even though many networks have multiple types of interactions or interactions that change over time, emphasis has traditionally been put on the study of simple or static networks. Recent work has however illustrated the advantages of exploiting the rich information of multislice networks: dynamic changes in modular organization of the human brain during learning are linked to the mastery of a novel skill (Mucha *et al.*, 2010; Bassett *et al.*, 2011); the multiple types of relations in human society help to understand the different roles humans play in different networks (Szell *et al.*, 2010); broad degree distributions enhance the vulnerability to failure of interdependent networks, in contrast to simple networks (Buldyrev *et al.*, 2010); and community detection in datasets with multiple types of similarities can be improved (Dong *et al.*, 2012).

Often, we want to study a signal *on* the network, not only the network itself. Wavelets are powerful tools in signal processing since they can localize signal contents in both space and frequency, and provide compact and multiscale representations of piecewise smooth signals. However, they are designed for signals defined on regular Euclidean spaces. Since real signals are often defined on

irregular domains multiple extensions to signals defined on graphs and manifolds have recently been proposed (Jansen *et al.*, 2009; Ram *et al.*, 2011; Hammond *et al.*, 2010; Coifman & Maggioni, 2006; Narang & Ortega, 2012). (We will use the terms network and graph interchangeably in this article.) These proposals can be split into spatial and spectral (frequency) designs. The former group includes Jansen *et al.* (2009) who based their method on lifting schemes, and Ram *et al.* (2011) who constructed wavelets for hierarchical trees. Hammond *et al.* (2010), Coifman & Maggioni (2006) and Narang & Ortega (2012) used spectral graph theory to construct spectral graph wavelets, “diffusion wavelets", and two-channel filter-banks for bipartite graphs, respectively. Importantly, graph wavelet transforms are adapted to the topology of the graph. Our proposed framework generalizes the spectral graph wavelet transform (SGWT) introduced in Hammond *et al.* (2010): a redundant wavelet transform for arbitrary weighted graphs that is based on the eigenspace of the graph Laplacian matrix and that can be efficiently applied to graphs with a large number of nodes by expressing the wavelets as iterates of the graph Laplacian operator.

Motivated by the promising results from multislice networks and the power of wavelets to analyze non-stationary signals, we sought to extend wavelets to multislice graphs. We expect multislice graph wavelets, which can adapt to the variable underlying graph topology, to be better suited to analyze signals defined on multislice graphs than “single slice" graph wavelets, constructed either separately for each slice or after simply averaging across them as is commonly done.

Our contributions in this paper are (1) to introduce a tight frame design for the SGWT, which is desirable because it preserves energy in the transformed domain and is self-reversible, which leads to a more efficient reconstruction scheme than in Hammond *et al.* (2010), (2) to make the connection between conventional and spectral graph wavelets explicit, and (3) to generalize the SGWT to multislice graphs. A multislice network, which consists of a set of networks, can be naturally represented with tensors (multi-way arrays). To reveal “hidden" structure in the variable graph topology, we propose to use tensor decompositions. Tensor decompositions are higher-order analogues of matrix factorizations, which are powerful tools, for, e.g. feature selection, but which are limited to two-dimensional data. To study data with three or more dimensions, tensor decompositions have been used in a wide variety of applications, such as chemometrics (Bro, 1997), psychoactivemetrics (Appellof & Davidson, 1981), face recognition (Vasilescu & Terzopoulos, 2002) and neuroscience (for a review see Kolda & Bader (2009)). The two most commonly used decompositions are the higher order singular value decomposition (HOSVD) and CANDECOMP/PARAFAC (CP), both of which are often considered as higher-order generalizations of the matrix singular value decomposition (SVD) and principal component analysis (PCA; Tucker, 1966; De Lathauwer *et al.*, 2000a; Carroll & Chang, 1970; Harshman, 1970). Each decomposition extends a different property of the SVD: The HOSVD decomposes a third-order tensor into three orthonormal subspaces associated with the different dimensions of a tensor, and CP decomposes it into a linear combination of rank-1 terms. We propose to use the HOSVD to learn “eigennetworks" from which we obtain novel networks by combining them in different ways. These networks and the resulting graph wavelets capitalize on the variability across a collection of networks.

In related work, Dong *et al.* (2012) introduced spectral clustering methods for multislice networks. The authors suggested that their methods may also be used to generalize other methods that are based on the eigenspace of the Laplacian matrix, but no applications other than clustering were shown. Their method differs from ours in that they derive one joint eigenspace for the multislice graph, whereas we derive multiple eigenspaces that capture the variability across the slices. We are not aware of any other

proposals to generalize wavelets to multislice graphs.

We apply the proposed method to different types of regular grid graphs and functional magnetic resonance imaging (fMRI) data to demonstrate that it can be a useful tool for analyzing signals defined on arbitrary weighted graphs with multiple types of interactions.

The paper is organized as follows. In Section 5.1.2, we briefly review the SGWT on which our framework is based. In Section 5.1.3 we introduce the construction of tight graph wavelet frames, rational dilation factors and the relationship between conventional and spectral graph wavelets. In Section 5.1.4, we propose a framework to design wavelets on multislice graphs and we present experimental results in Section 5.1.5.

Notations

Scalars are denoted by lowercase letters (e.g., a), vectors by boldface lowercase letters (e.g., $\mathbf{a}$), matrices by boldface capital letters (e.g., $\mathbf{A}$), and tensors by Euler script letters (e.g., $\mathcal{A}$).

The order of a tensor is its number of dimensions, also known as modes. For third-order tensors, slices are two-dimensional subarrays obtained by fixing one index: $\mathbf{A}_{i::}$, $\mathbf{A}_{:j:}$ and $\mathbf{A}_{::k}$ for a horizontal, vertical and frontal slice, respectively.

The multiplication of a tensor with a matrix in mode n is denoted by $\mathcal{A} \times_n \mathbf{B}$. For a third-order tensor $\mathcal{A} \in \mathbb{R}^{I \times J \times K}$, its mode-1 product with the matrix $\mathbf{B} \in \mathbb{R}^{M \times I}$ is a tensor of size $(M \times J \times K)$ with entries

$$(\mathcal{A} \times_1 \mathbf{B})_{mjk} = \sum_{i=1}^{I} a_{ijk} b_{mi}. \tag{5.1}$$

Using the n-mode product notation, the matrix product $\mathbf{G} = \mathbf{U} \cdot \mathbf{F} \cdot \mathbf{V}^T$ can be rewritten as $\mathbf{G} = \mathbf{F} \times_1 \mathbf{U} \times_2 \mathbf{V}$.

A tensor can be converted to a matrix by unfolding it along any mode. The mode-1 matricization or unfolding of a third-order tensor $\mathcal{A}$ is the $I \times JK$ matrix $\mathbf{A}_{(1)}$ with entries $(\mathbf{A}_{(1)})_{im} = l_{ijk}$ where $m = (j-1)K + k$, i.e., $\mathbf{A}_{(1)}$ is a reordered concatenation of the frontal slices of the tensor.

5.1.2 Review of spectral graph wavelets

We start with a short overview of the construction of the SGWT and refer to (Hammond *et al.*, 2010) for details and proofs.

Classical one-dimensional wavelets

To solve the problem of how to define shifting and scaling of wavelets on irregular domains, Hammond et al. appealed to the Fourier domain. The one-dimensional continuous wavelet $\psi_{s,a}(x) = \frac{1}{s}\psi(\frac{x-a}{s})$, at scale s and location a, can be defined in the Fourier domain as

$$\psi_{s,a}(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{\psi}(s\omega) e^{-j\omega a} e^{j\omega x} d\omega. \tag{5.2}$$

For ease of understanding of the SGWT design that will follow, we highlight several properties of Eq. (5.2):

- the wavelet is represented on complex exponentials, which are eigenfunctions of the one-dimensional Laplacian operator; i.e., $\frac{d^2}{dx^2} e^{j\omega x} = -\omega^2 e^{j\omega x}$;
- shifting the wavelet to location a corresponds to a multiplication by $e^{-j\omega a}$;
- the wavelet can be interpreted as scaled bandpass filter and scaling ψ by $1/s$ corresponds to scaling $\hat{\psi}$ with s.

Shifting and scaling of the classical wavelet can thus be defined in the Fourier domain. To generalize wavelets to graph, we will need the analogue of the Fourier domain for graphs.

Weighted graphs

Let $\mathcal{G} = (V, E, W)$ be an undirected graph consisting of $|V| = N$ nodes that are connected by the set of edges E with nonnegative weights W. The $N \times N$ adjacency matrix $\mathbf{A}$ for a graph without loops is then given by the non-diagonal entries

$$a_{ij} = \begin{cases} w(i,j), & \text{if } (i,j) \in E, \\ 0, & \text{otherwise.} \end{cases} \tag{5.3}$$

The diagonal degree matrix $\mathbf{D}$ is given by $d_{ii} = \sum_j a_{ij}$ and the Laplacian matrix $\mathbf{L}$ by $\mathbf{L} = \mathbf{D} - \mathbf{A}$. The normalized version of the Laplacian matrix $\mathbf{L}$ is defined as $\widetilde{\mathbf{L}} = \mathbf{D}^{-\frac{1}{2}} \mathbf{L} \mathbf{D}^{-\frac{1}{2}}$.

The eigenspaces of the Laplacian matrices form the graph analogue of the Fourier domain. $\mathbf{L}$ and $\widetilde{\mathbf{L}}$ are symmetric, positive semi-definite matrices and can thus be decomposed into non-negative eigenvalues and eigenvectors. For $\mathbf{L}$ the decomposition is given by

$$\mathbf{L} = \mathbf{X} \Lambda \mathbf{X}^T, \tag{5.4}$$

where $\mathbf{X} = [\chi_1 | \chi_2 | ... | \chi_N]$ is an orthonormal matrix containing N eigenvectors and Λ is a diagonal matrix whose entries equal the associated eigenvalues $\lambda_1 = 0 \leq \lambda_1 \leq ... \leq \lambda_N$ (Chung, 1997). For $\widetilde{\mathbf{L}}$, $\lambda_\ell \in [0, 2]$ with $\ell = 1, 2, ...N$, and the eigenvectors are different from those of $\mathbf{L}$ (except if $\widetilde{\mathbf{L}} = \mathbf{L}$, of course).

Spectral graph wavelets

In analogy to the classical wavelet of 5.1.2, we can now define the spectral graph wavelet $\psi_{t,n}(m)$, at scale t and node n, in the eigenspace of the graph Laplacian matrix $\mathbf{L}$ or $\widetilde{\mathbf{L}}$:

$$\psi_{t,n}(m) = \sum_{\ell=0}^{N-1} g(t\lambda_\ell) \chi_\ell^*(n) \chi_\ell(m), \tag{5.5}$$

where g is the wavelet generating function defined in the spectral graph domain. We now compare Eq. (5.2) and Eq. (5.5):

- the wavelet is represented on the eigenvectors $\chi_\ell(m)$ of the Laplacian matrix, replacing $e^{i\omega x}$;
- shifting the wavelet to node n corresponds to a multiplication by $\chi_\ell^*(n)$, replacing $e^{-i\omega a}$;
- g behaves as a scaled bandpass filter, replacing $\hat{\psi}$;
- the role of the frequency ω is played by the eigenvalue λ_ℓ.

Note that in practice the continuous scaling parameter t will be sampled to a finite number of scales; i.e., we have $\{t_j\}_{j=1}^{J}$ for J scales.

We also need to define the scaling function $\phi_n(m)$ to capture the residual "low-pass" components:

$$\phi_n(m) = \sum_{l=0}^{N-1} h(t_J \lambda_\ell) \chi_\ell^*(n) \chi_\ell(m), \tag{5.6}$$

where h is the scaling function generator. An example of exact specification of $g(\cdot)$ and $h(\cdot)$ was given in Hammond *et al.* (2010).

As in the one-dimensional setting, the wavelet and scaling coefficients are given by the inner product of the graph signal f with the wavelets $\psi_{t,n}$ and the scaling function ϕ_n, respectively:

$$W_f(t,n) = \langle \psi_{t,n}, f \rangle = \sum_{m=1}^{N} \psi_{t,n}^*(m) f(m), \tag{5.7}$$

$$S_f(n) = \langle \phi_n, f \rangle = \sum_{m=1}^{N} \phi_n^*(m) f(m). \tag{5.8}$$

Because it is infeasible to explicitly compute the eigenspace of the Laplacian matrix for large graphs, the authors proposed a fast method to compute the forward transform. The method is based on approximating the kernels g and h with a low-dimensional Chebyshev polynomial:

$$g(t_j x) = \frac{1}{2} c_{j,0} + \sum_{k=1}^{M_j} c_{j,k} T_a^k(x), \tag{5.9}$$

where $c_{j,k}$ denotes the Chebyshev coefficients, M_j the order of the polynomial approximation, and $T_a^k(x)$ the shifted Chebyshev polynomials $T_a^k(x) = T^k(\frac{x-a}{a})$ of order k. The fast approximate spectral graph wavelet forward transform is given by

$$\widetilde{W}_f(t_j, n) = \frac{1}{2} c_{j,0} f + \sum_{k=1}^{M_j} c_{j,k} T_a^k(\mathbf{L}) f, \tag{5.10}$$

$$\widetilde{S}_f(t_j, n) = \frac{1}{2} c_{0,0} f + \sum_{k=1}^{M_0} c_{0,k} T_a^k(\mathbf{L}) f. \tag{5.11}$$

To recover a graph signal f from a given set of wavelet coefficients $c = \widetilde{W} f$, where c includes the wavelet and scaling coefficients $\widetilde{W}_f$ and $\widetilde{S}_f$, and $\widetilde{W} : \mathbb{R}^N \to \mathbb{R}^{N(J+1)}$ represents the approximate SGWT, we use a left-inverse M s.t. $MWf = f$. Since the SGWT is a redundant transform, there is an infinite number of possible inverses. Hammond *et al.* (2010) chose to use the pseudoinverse that solves $(\widetilde{W}^* \widetilde{W}) f = \widetilde{W}^* c$, where $\widetilde{W}^*$ is the adjoint operator. For large graphs, the pseudoinverse can be calculated by an iterative conjugate gradient method.

5.1.3 Tight frame graph wavelets

Tight frame design

Hammond *et al.* (2010) highlighted the need to optimize the design of the wavelet generating function $g(\cdot)$ and scaling function $h(\cdot)$ for specific applications. Here, we are interested in constructions that lead to tight (or Parseval) wavelet frames because of their important property of energy conservation between the original and transformed domain and the easy signal reconstruction (Benedetto & Fickus, 2003).

Definition 1 *A family of wavelets* $\{\psi_{t,n}\}_{t\in J, n\in N}$ *forms a frame of* $\ell_2(V)$*, if there exist frame bounds* $A, B > 0$ *such that*

$$\forall f \in \ell_2(V), \quad A||f||^2 \leq \sum_t \sum_n |\langle f, \psi_{t,n}\rangle|^2 \leq B||f||^2. \tag{5.12}$$

If $A = B = 1$*, the family* $\{\psi_{t,n}\}$ *forms a Parseval frame and preserves energy:*

$$\forall f \in \ell_2(V), \quad ||f||^2 = \sum_t \sum_n |\langle f, \psi_{t,n}\rangle|^2. \tag{5.13}$$

A Parseval frame is the normalized version of a tight frame ($A = B$). Notably, Parseval frames have a simple reconstruction formula where the analysis functions are used at the synthesis side:

$$\forall f \in \ell_2(V), \quad f = \sum_t \sum_n \langle f, \psi_{t,n}\rangle \psi_{t,n} \tag{5.14}$$

Thus, the fast approximate SGWT can be used for both the analysis and synthesis steps, avoiding the iterative computation of the pseudoinverse. Note that the proposed analysis/synthesis scheme uses an approximation for both steps, whereas the pseudoinverse proposed in Hammond *et al.* (2010) compensates for approximation errors made in the analysis step. Therefore, the polynomial order M_j needs to be chosen high enough in the former case to ensure perfect reconstruction.

We propose to generate Parseval graph wavelet frames by designing the wavelet generating kernel g and scaling function h in the spectral graph domain in analogy to classical definitions in the Fourier frequency domain (Table 5.1, Fig. 5.3b). The J dyadic spatial scales are defined as $t_j = \frac{a}{\lambda_{max}} 2^j$ for $j = 1, \ldots, J$ and the factor $\frac{a}{\lambda_{max}}$ is chosen to ensure full spectral coverage of the kernels.

All of these constructions are bandlimited functions, and the union of the functions $|g(t_j\lambda)|^2$ and $|h(t_j\lambda)|^2$ forms a partition of unity: $\sum_{j=1}^{J} |g(t_j\lambda)|^2 + |h(t_J\lambda)|^2 = 1$.

Rational dilation factors

The eigenvalues λ_ℓ can be unevenly distributed depending on the nature of the graph (see Fig. 5.3a for an example). For increased flexibility in defining the wavelet subbands, we also provide rational frame designs. Extensions to the rational case of Meyer and Shannon constructions for classical wavelets were originally proposed by Auscher (1992) and have further been generalized by Baussard *et al.* (2004).

The rational wavelet generating kernel $g(\cdot)$ and scaling function $h(\cdot)$ with rational dilation factor

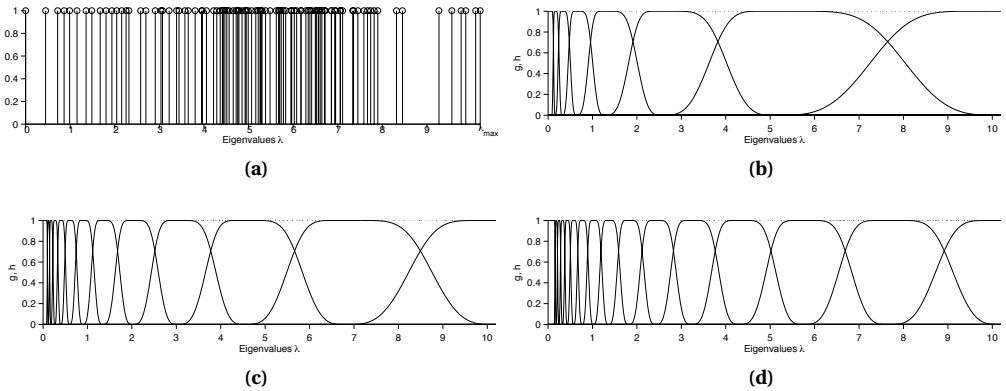


Figure 5.3 – (a) Eigenvalues of the Swiss Roll graph from Hammond *et al.* (2010) with $N = 50$ nodes, and $\sigma = 0.3$. Scaling function h, wavelet generating kernels g, and sum of squares (dotted line) of a Meyer-like rational wavelet frame with $M = 2$ (b), $\frac{3}{2}$ (c) and $\frac{4}{3}$ (d) and $\epsilon = (1+M)^{-1}$.

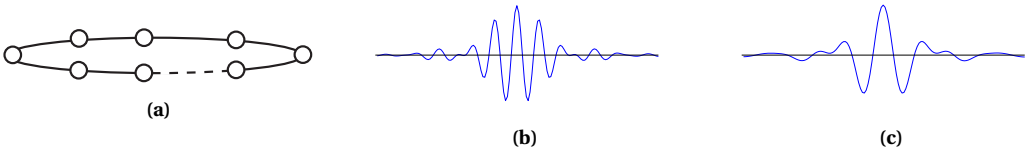


Figure 5.4 – (a) Schematic of a cycle graph that corresponds to an equispaced sampling grid of a one-dimensional signal; (b) the spectral graph wavelet for the Meyer generating window $g(\lambda_\ell)$ does not equal the Meyer wavelet, (c) but it does for the generating window $g(\omega(\lambda_\ell))$ with $\omega = \arccos(1 - \lambda_\ell/2)$.

$M = \frac{q+1}{q}$, $q \in \mathbb{Z}$ are defined as follows for a Meyer-like wavelet frame:

$$g(\lambda) = \begin{cases} \sin\left(\frac{\pi}{2}\nu\left(q\left(\frac{\lambda}{a}-1\right)\right)\right) & \forall \lambda \in]a, Ma] \\ \cos\left(\frac{\pi}{2}\nu\left(q\left(\frac{\lambda}{Ma}-1\right)\right)\right) & \forall \lambda \in]Ma, M^2a] \\ 0 & \text{elsewhere} \end{cases} \tag{5.15}$$

$$h(\lambda) = \begin{cases} 1 & \forall \lambda \in [0, a] \\ \cos\left(\frac{\pi}{2}\nu\left(q\left(\frac{\lambda}{a}-1\right)\right)\right) & \forall \lambda \in]a, Ma] \\ 0 & \text{elsewhere} \end{cases} \tag{5.16}$$

where $a = (q-\epsilon)b$, $\epsilon =]0, (1+M)^{-1}]$ and $a, b \in \mathbb{R}^+$. Clearly, the classical dyadic dilation corresponds to $M = 2$. The J wavelet scales are defined as $t_j = \frac{a}{\lambda_{max}} M^j$ for $j = 1, \ldots, J$. Figs. c and (d) show the spectra of two Meyer-like rational wavelets.

Conventional and spectral graph wavelets

As stated in 5.1.2, the eigenvalues λ_ℓ of the graph Laplacian play the role of frequencies ω. However, the exact relation between spectral and conventional Fourier frequencies was not further deepened

Table 5.1 – Specification of wavelet and scaling generating functions that lead to Parseval wavelet frames ($a \in \mathbb{R}^+$)

Wavelet type	Spectral frequency response
Meyer (Meyer, 1986)	$g(\lambda) = \begin{cases} \sin\left(\frac{\pi}{2}\nu\left(\frac{\lambda}{a}-1\right)\right) & \forall\lambda\in]a,2a] \\ \cos\left(\frac{\pi}{2}\nu\left(\frac{\lambda}{2a}-1\right)\right) & \forall\lambda\in]2a,4a] \\ 0 & \text{elsewhere} \end{cases}$ $h(\lambda) = \begin{cases} 1 & \forall\lambda\in[0,a] \\ \cos\left(\frac{\pi}{2}\nu\left(\frac{\lambda}{a}-1\right)\right) & \forall\lambda\in]a,2a] \\ 0 & \text{elsewhere} \end{cases}$ $\nu(x) = x^4(35-84x+70x^2-20x^3)$
Simoncelli (Portilla & Simoncelli, 2000)	$g(\lambda) = \begin{cases} \cos\left(\frac{\pi}{2}\log_2\left(\frac{\lambda}{2a}\right)\right) & \forall\lambda\in]a,4a] \\ 0 & \text{elsewhere} \end{cases}$ $h(\lambda) = \begin{cases} 1 & \forall\lambda\in[0,a] \\ \cos\left(\frac{\pi}{2}\log_2\left(\frac{\lambda}{a}\right)\right) & \forall\lambda\in]a,2a] \\ 0 & \text{elsewhere} \end{cases}$
Shannon	$g(\lambda) = \begin{cases} 1 & \forall\lambda\in]a,2a] \\ 0 & \text{elsewhere} \end{cases}$ $h(\lambda) = \begin{cases} 1 & \forall\lambda\in[0,a] \\ 0 & \text{elsewhere} \end{cases}$
Papadakis (Romero *et al.*, 2009)	$g(\lambda) = \begin{cases} \sqrt{\frac{1+\sin\left(\frac{3\pi\lambda}{2a}\right)}{2}} & \forall\lambda\in]a,\frac{5}{3}a] \\ 1 & \forall\lambda\in]\frac{5}{3}a,2a] \\ \sqrt{\frac{1-\sin\left(\frac{3\pi\lambda}{4a}\right)}{2}} & \forall\lambda\in]2a,\frac{10}{3}a] \\ 0 & \text{elsewhere} \end{cases}$ $h(\lambda) = \begin{cases} 1 & \forall\lambda\in[0,a] \\ \sqrt{\frac{1-\sin\left(\frac{3\pi\lambda}{2a}\right)}{2}} & \forall\lambda\in]a,\frac{5}{3}a] \\ 0 & \text{elsewhere} \end{cases}$
Held (Held *et al.*, 2010)	$g(\lambda) = \begin{cases} \cos\left(2\pi\mu\left(\frac{\lambda}{8a}\right)\right) & \forall\lambda\in]a,2a] \\ \sin\left(2\pi\mu\left(\frac{\lambda}{16a}\right)\right) & \forall\lambda\in]2a,4a] \\ 0 & \text{elsewhere} \end{cases}$ $h(\lambda) = \begin{cases} 1 & \forall\lambda\in[0,a] \\ \sin\left(2\pi\mu\left(\frac{\lambda}{8a}\right)\right) & \forall\lambda\in]a,2a] \\ 0 & \text{elsewhere} \end{cases}$ $\mu(x) = -1+24x-144x^2+256x^3$

in Hammond *et al.* (2010). Here, we want to find this relation. As we have introduced wavelet generating functions from previously proposed wavelet bases, we want to define the mapping $w : \lambda_\ell \to \omega_\ell$ that leads to the classical Meyer wavelet when using the Meyer generating function and a Cartesian grid. We use a cycle graph because it corresponds to an equispaced sampling grid of a one-dimensional signal with periodic boundary conditions, and the eigenvectors of its Laplacian matrix are the basis vectors of the discrete Fourier transform (Fig. 5.4a; Strang, 1999). Without any mapping w, the spectral graph

wavelet does not yield the classical Meyer wavelet (Fig. 5.4b). Intuitively, one would expect a mapping $\lambda_\ell = \omega_\ell^2$ since the Laplacian corresponds to a second-order derivative operator ($\frac{d^2 e^{j\omega x}}{dx^2} = -\omega^2 e^{j\omega x}$), but the relation is slightly more subtle. To understand it, we consider the eigenvalues of the cycle graph (Chung, 1997):

$$\lambda_\ell = 2(1-\cos(2\pi\ell/N)) \quad \in [0,4] \tag{5.17}$$

with $\ell = 0,1,...,N-1$; N the number of nodes, 2 the degree of the nodes and $2\pi\ell/N = \omega_\ell$ the conventional Fourier frequency. Inversion of Eq. (5.17) leads to the mapping $w : \lambda_\ell \rightarrow \omega_\ell$

$$\omega_\ell = \arccos(1-\lambda_\ell/2) \quad \in [0,\pi], \tag{5.18}$$

which we incorporate in the wavelet generating function as $g(w(\cdot))$. This mapping linearizes the eigenvalues and the spectral graph wavelet yields the classical Meyer wavelet (Fig. 5.4(c)). Note that the behavior of arccos around the origin is quadratic, as expected, but that the eigenvalues at the upper part of the Laplacian spectrum are saturated.

Because the spectrum of any graph Laplacian matrix $\mathbf{L}$ or $\widetilde{\mathbf{L}}$ is contained in $[0, 2d_{\max}]$, we will incorporate the mapping

$$\omega_\ell = \arccos(1-\lambda_\ell/d_{\max}) \tag{5.19}$$

in the wavelet generating function g and scaling function generator h.

Interestingly, the same relation between spectral and conventional Fourier frequencies has been reported when considering the wave equation on graphs. For the interested reader, we further discuss this result in the appendix.

5.1.4 Wavelets on multislice graphs

We propose to construct wavelets on multislice graphs by first estimating eigennetworks of a multislice graph, which capture the variability across slices (5.1.4), and then, within this eigennetwork construction, apply the SGWT (5.1.4).

Eigenspace of multislice gaphs

Dynamic networks or networks connected by different types of edges can be represented by a multislice network, where each slice corresponds to one time point or type of interaction (Fig. 5.5). A multislice graph with N nodes and K different types of edges is described by a third-order adjacency tensor $\mathcal{A} \in \mathbb{R}^{N\times N\times K}$ with the individual adjacency matrices $\mathbf{A}_k \in \mathbb{R}^{N\times N}$ as its frontal slices: $\mathbf{A}_{::k} = \mathbf{A}_k$, for $k = 1,...,K$.

Here, we use tools from multilinear algebra, the HOSVD in particular, to extract orthonormal subspaces from the adjacency tensor, from which “novel" graphs can be constructed that capture the variability across the edge measures.

Proposition 1 *The HOSVD of a tensor $\mathcal{A} \in \mathbb{R}^{N\times N\times K}$ with symmetric frontal slices ($\mathbf{A}_k = \mathbf{A}_k^T$ for all $k =$*

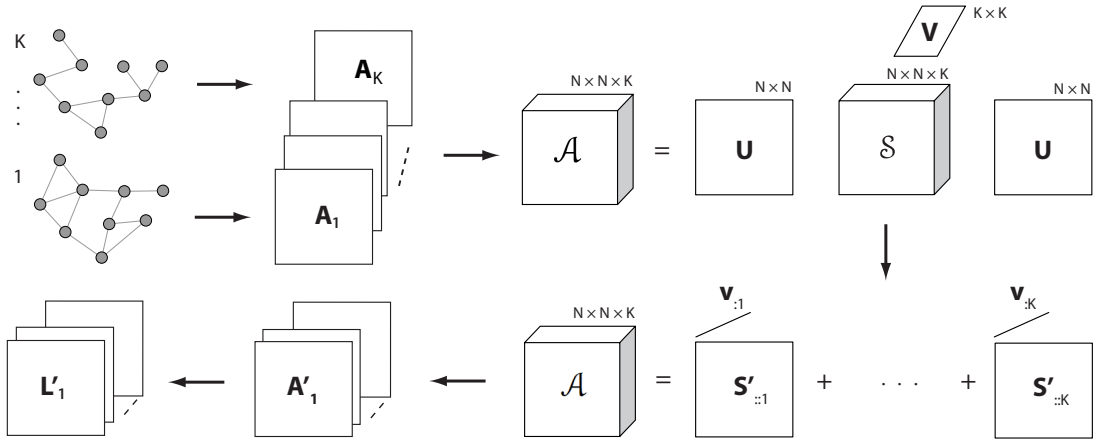


Figure 5.5 – An adjacency tensor $\mathcal{A}$ is constructed from a stack of K adjacency matrices $\mathbf{A}_k$. The HOSVD decomposes $\mathcal{A}$ into a core tensor $\mathcal{S}$ with symmetric frontal slices and orthonormal matrices $\mathbf{U}$ and $\mathbf{V}$. The decomposition can also be interpreted as a weighted sum of eigennetworks $\mathbf{S'}_{::k} = \mathbf{U}\mathbf{S}_{::k}\mathbf{U}^T$. New adjacency matrices $\mathbf{A'}_m$ are built from linear combinations of the eigennetworks, which each reflect different properties of the underlying network. The adjacency matrices are then converted to Laplacian matrices $\mathbf{L'}_m$ to apply the SGWT.

$1,...,K$) *yields the decomposition*

$$\mathcal{A} = \mathcal{S} \times_1 \mathbf{U} \times_2 \mathbf{U} \times_3 \mathbf{V} \tag{5.20}$$

where $\mathcal{S} \in \mathbb{R}^{N\times N\times K}$ *is called the core tensor and* $\mathbf{U} \in \mathbb{R}^{N\times N}$ *and* $\mathbf{V} \in \mathbb{R}^{K\times K}$ *are orthonormal mode matrices (Fig. 5.5). The* 1*-mode singular values are defined by* $\sigma_i^{(1)} = \sqrt{\langle \mathbf{S}_{1::}, \mathbf{S}_{1::} \rangle}$ *for* $i = 1,..,N$*, and are ordered* $\sigma_1^{(1)} \geq \sigma_2^{(1)} \geq ... \geq \sigma_N^{(1)}$*. The singular values for the other two modes are analogous.*

The core tensor $\mathcal{S}$ governs the interactions between the factors represented in the mode matrices; i.e., between the nodes and various edge measures, and the mode matrix $\mathbf{U}$ spans a joint subspace of all adjacency matrices.

Definition 2 *The core tensor* $\mathcal{S}$ *can be transformed into an "eigennetwork" tensor* $\mathcal{S}'$ *by the joint subspace* $\mathbf{U}$*:*

$$\mathcal{S}' = \mathcal{S} \times_1 \mathbf{U} \times_2 \mathbf{U} \tag{5.21}$$

with $\mathcal{S}' \in \mathbb{R}^{N\times N\times K}$*. The individual eigennetworks* $\mathbf{S'}_{::k}$ *are given by*

$$\mathbf{S'}_{::k} = \mathbf{S}_{::k} \times_1 \mathbf{U} \times_2 \mathbf{U} = \mathbf{U}\mathbf{S}_{::k}\mathbf{U}^T \tag{5.22}$$

with $k = 1,...,K$ *and using the slice-wise representation of Harshman & Hong (2002). The eigennetworks are orthogonal in the sense of the scalar product* $\langle \mathbf{S'}_{::i}, \mathbf{S'}_{::j} \rangle = 0$*, for* $i \neq j$*, and symmetric since* $\mathcal{S}$ *shares the symmetry of* $\mathcal{A}$ *(De Lathauwer* et al.*, 2000a):* $\mathbf{S'}_{::k}{}^T = \left(\mathbf{U}\mathbf{S}_{::k}\mathbf{U}^T\right)^T = \mathbf{U}\mathbf{S}_{::k}{}^T\mathbf{U}^T = \mathbf{U}\mathbf{S}_{::k}\mathbf{U}^T = \mathbf{S'}_{::k}$.

The first eigennetwork $\mathbf{S'}_{::1}$ is the average over all networks and the additional $K-1$ eigennetworks $\mathbf{S'}_{::k}$ capture the variability across them. The idea of extracting eigennetworks from a tensor using

HOSVD is similar to work from Vasilescu & Terzopoulos (2002) who proposed “TensorFaces", a facial image recognition technique.

The adjacency tensor $\mathcal{A} = \mathcal{S}' \times_3 \mathbf{V}$ can now also be interpreted as a weighted sum of eigennetworks, as illustrated in Fig. 5.5 (Eldén, 2007):

$$\mathbf{A}_{::k} = \sum_{t}^{T} v_{kt} \mathbf{S}'_{::t}. \tag{5.23}$$

For a graph with a large number of nodes N, it is infeasible to compute the eigennetworks $\mathbf{S}'_{::k}$. However, noting that $\mathcal{S}' = \mathcal{A} \times_3 \mathbf{V}^T$, since $(\mathcal{S}' \times_3 \mathbf{V}) \times_3 \mathbf{V}^T = \mathcal{S}' \times_3 (\mathbf{V}\mathbf{V}^T) = \mathcal{S}'$ (De Lathauwer *et al.*, 2000a), we can efficiently compute them in an indirect way if $K << N$:

$$\mathbf{S}'_{::k} = \sum_{t}^{T} v_{tk} \mathbf{A}_{::t}, \tag{5.24}$$

which is a weighted sum of the frontal slices of $\mathcal{A}$. $\mathbf{V}$ can be directly computed as the left singular matrix of the mode-3 matrix unfolding of $\mathcal{A}$: $\mathbf{A}_{(3)} = \mathbf{V}\Sigma\mathbf{W}^T$. For large graphs it is more efficient to compute $\mathbf{V}$ by decomposing $\mathbf{A}_{(3)}\mathbf{A}_{(3)}^T = \mathbf{V}\Sigma^2\mathbf{V}^T$.

Graph wavelet design

In many real-world cases, most of the variability in the edge measures can be reasonably assumed to be captured by the first few eigennetworks with the largest singular values. We can thus learn a multilinear subspace of our original adjacency tensor by keeping only the eigennetworks with the largest singular values. Note that the tensor $\hat{\mathcal{A}}$ obtained by discarding the smallest 3-mode singular values is in general not the optimal rank-N, N, κ ($\kappa < K$) approximation of a tensor $\mathcal{A}$, because in HOSVD the mode matrices are optimized for each mode separately, neglecting interactions among them. However, the truncated HOSVD is a good approximation with the error bounded by $||\mathcal{A} - \hat{\mathcal{A}}||^2 \leq \sum_{i=\kappa+1}^{K} (\sigma_i^{(3)})^2$ and is much faster than the otherwise required iterative algorithm (De Lathauwer *et al.*, 2000b).

We propose to design adjacency matrices $\mathbf{A'}_m$, $m = 1, ..., M$, that capture different properties of the underlying networks by combining the eigennetworks:

$$\mathbf{A'}_m = \sum_{k=1}^{\kappa} \alpha_k^{(m)} \mathbf{S}'_{::k} \quad \text{s.t. } \mathbf{A}'_m \geq \mathbf{0} \tag{5.25}$$

where $\alpha_k^{(m)} \in \mathbb{R}$. The SGWT is only defined for graphs with nonnegative weights and the set of $\alpha_k^{(m)}$ must therefore be chosen such that $\mathbf{A'}_m \geq \mathbf{0}$. Since $\mathbf{S}'_{::1}$ is the average graph, or “backbone" of the multislice graph, and only all-postiive or all-negative eigennetwork, $\alpha_1^{(m)}$ must be $\neq 0$.

The weights $\alpha_k^{(m)}$ can also be obtained by a constrained least-squares fitting problem, which exploits the weights in $\mathbf{V}$ associated with each eigennetwork k:

$$\min_{\alpha_k^{(m)}} \left\| \sum_{k=1}^{\kappa} \alpha_k^{(m)} \mathbf{v}_{:k} - \mathbf{g} \right\|^2 \quad \text{s.t. } \mathbf{A}'_m \geq \mathbf{0} \tag{5.26}$$

where $\mathbf{g}$ is a pre-defined variation between the slices of the network that we like $\mathbf{A}'_m$ to capture. This strategy will not be further explored here.

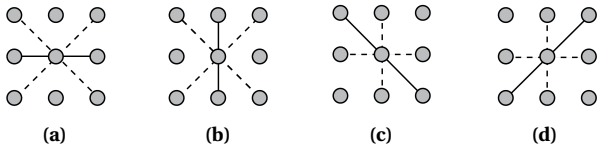


Figure 5.6 – Neighborhoods of 4 different grid graphs. Dashed lines indicate edge weights that were penalized by a factor 1/2 so as to enhance one direction: (a) horizontal, (b) vertical, (c) diagonal top left to bottom right, (d) diagonal bottom left to top right.

After calculating $\mathbf{L}'_m$, we can then use the SGWT or its approximation for large graphs to construct wavelets.

5.1.5 Experimental results and discussion

Graph wavelets on multiplex graphs

An image can be interpreted as a grid graph where each pixel is a node that can be connected to neighboring pixels/nodes according to chosen properties. To illustrate the potential of our design to extract meaningful information from multislice graphs, we constructed a tensor from 4 different grid graphs. Each node was connected to at most 6 neighbors along different directions; i.e., two neighboring nodes were disconnected if their intensity values in the image differed by more than 40. The image was pre-smoothed prior to estimating the adjacency matrix to better approximate the gradient operator (Gaussian kernel, $\sigma = 1$; Di Zenzo, 1986). The slices of the adjacency tensor $\mathcal{A}$ were ordered as shown in Fig. 5.6.

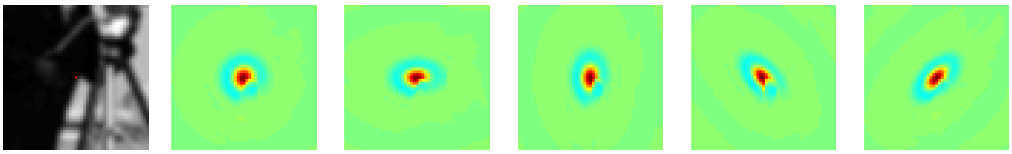
We obtained 4 eigennetworks and associated weights $\mathbf{V}$ from the HOSVD of $\mathcal{A}$. The columns of $\mathbf{V}$ capture the variability across mode 3 of $\mathcal{A}$ (i.e., the direction of the edges):

$$\begin{bmatrix} -0.47 & 0.73 & 0 & 0.50 \\ -0.49 & -0.69 & 0 & 0.54 \\ -0.52 & -0.01 & 0.71 & -0.47 \\ -0.52 & -0.01 & -0.71 & -0.49 \end{bmatrix}$$

with associated singular values $\sigma_1^{(3)} = 97, \sigma_2^{(3)} = 48, \sigma_3^{(3)} = 31, \sigma_4^{(3)} = 18$. We observe that the first column is the average of all slices, the second one contrasts the horizontal and vertical edges (large values in row 1 and 2), the third one the diagonal ones, and the last one rectangular and diagonal ones.

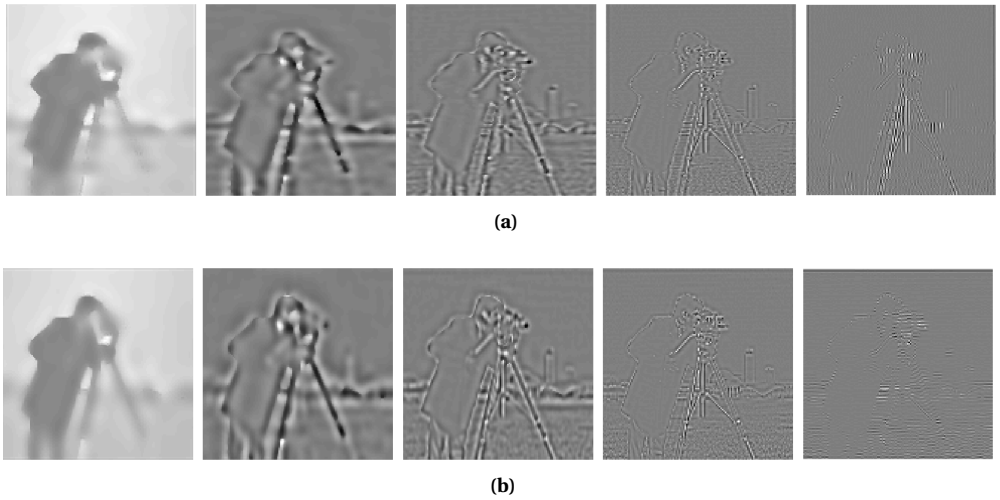
In Fig. 5.7 we show 5 "directional" graph wavelets that were implemented by combining the eigennetworks in different ways. Whereas wavelets constructed using the average graph diffuse homogeneously (i.e. there is no directional information, Fig. 5.7b), they diffuse preferentially along a certain direction when the average graph is combined with one of the other 3 eigennetworks, making them sensitive to different types of orientations in an image (Figs. 5.7c-f). Since neighboring nodes were only connected if their intensity values were sufficiently close, the wavelets diffuse along underlying boundaries, such as the coat.

Fig. 5.8 shows two graph wavelet decompositions of the cameraman image of size 128×128, where the SGWT was built from "horizontal" $\mathbf{A}'_2$ and "vertical" $\mathbf{A}'_3$ (wavelets shown in Figs. 5.7c and d). The energy of the wavelet coefficients is high for different types of features of the image: At the finest scale,



(a) (b) (c) (d) (e) (f)

Figure 5.7 – Graph wavelets constructed from various combinations of eigennetworks diffuse along different directions: (a) location of the wavelet on the smoothed cameraman image, (b) $\mathbf{A'}_1 = -\mathbf{S'}_{::1}$: isotropic for the average graph, (c) $\mathbf{A'}_2 = -0.5\mathbf{S'}_{::1} + 0.7\mathbf{S'}_{::2}$: horizontal, (d) $\mathbf{A'}_3 = -0.5\mathbf{S'}_{::1} - 0.7\mathbf{S'}_{::2}$: vertical, (e) $\mathbf{A'}_4 = -0.5\mathbf{S'}_{::1} + 0.8\mathbf{S'}_{::3}$: diagonal, (f) $\mathbf{A'}_5 = -0.5\mathbf{S'}_{::1} - 0.8\mathbf{S'}_{::3}$: diagonal. The first frontal slice $\mathbf{S'}_{::1}$ was multiplied by $-0.5 = 1/\sum_i \mathbf{v}_{i1}$, which turns it into the average graph. Wavelets are shown at a coarse scale.



(a)

(b)

Figure 5.8 – Wavelet coefficients of the cameraman image (128 × 128 pixels), where the graph wavelet transforms were built from (a) $\mathbf{A'}_2$, (b) $\mathbf{A'}_3$. $J = 4$ scales, Simoncelli-like design, and scaling and coarse to fine scale wavelet coefficients from left to right.

which detects edges not expected by the underlying graph topology, the energy for the “horizontal" transform is high around vertical edges, such as the tripod, and for the “vertical" transform around horizontal edges, such as the camera.

For a graph corresponding to a Cartesian grid and when using the mapping introduced in 5.1.3, we know that the spectral graph wavelets are equivalent to classical wavelets. The advantage of using graph wavelets, however, is that they are not restricted to Cartesian grids. Weighted graphs provide a flexible model to describe irregular data domains and the wavelets can thus be adapted to many different applications. Here, we have used this flexibility to incorporate knowledge about directionality and underlying boundaries. With the proposed approach for multislice graphs, we also not only have access to an average graph ($\mathbf{S'}_{::1}$), which corresponds to the common way of handling multislice graphs, but several eigennetworks that capture the variability across the direction of the edges.

Graph wavelets on dynamic graphs

Cognitive function and dynamic adaptation emerge through the integrated activity of many brain regions. Using functional magnetic resonance imaging (fMRI), functionally connected brain regions

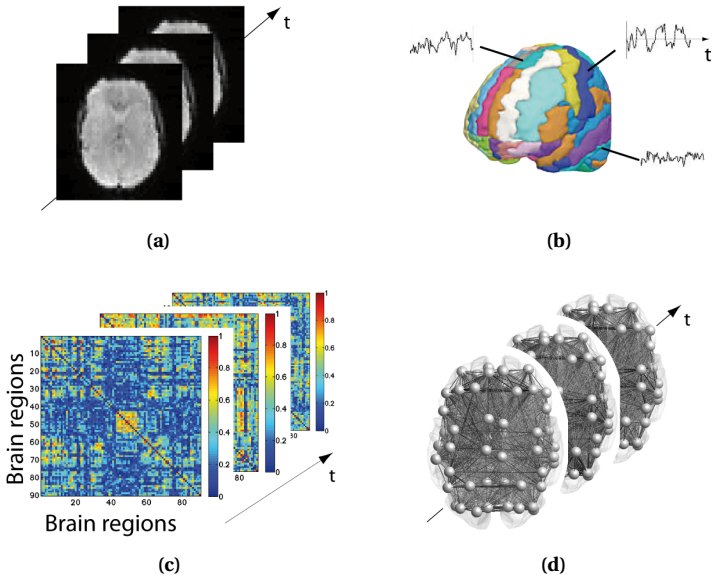


Figure 5.9 – fMRI analysis pipeline: (a) raw fMRI data; (b) parcellation of the brain into regions (network nodes), and estimation of their regional spatially-averaged activity; (c) to estimate changes in network architecture over time, we computed the temporal correlation between the regional activity, averaged across subjects, of all pairs of regions using a sliding window approach (weighted network edges); (d) and obtained $K = 125$ functional brain networks.

are identified as those exhibiting temporal dependence in their activity patterns. Studies representing the brain as a static, functionally connected network have already led to unique insights into the organization of the brain, but understanding how the architecture of the brain varies with cognitive function or across subjects is a crucial frontier (Sporns, 2010; Bassett & Gazzaniga, 2011). Chang & Glover (2010) were among the first to highlight that functional connectivity is not stationary throughout an fMRI scanning session (with a focus on one brain region), and Bassett *et al.* (2011) recently studied dynamic reconfigurations in the modular organization of brain networks during learning.

We explored the potential of our proposed framework to capture dynamic changes of the human brain network in a previously studied dataset (Eryilmaz *et al.*, 2011; Richiardi *et al.*, 2011). It was shown that average functional connectivity differs between a “resting" and “movie watching" cognitive state and enough so that a subject’s cognitive state could be predicted using a classifier algorithm (Richiardi *et al.*, 2011).

The fMRI data was acquired from 15 healthy subjects during a simple block design, which consisted of alternating cycles of watching a short movie excerpt (40-60 sec), followed by a period where subjects were instructed to rest with eyes closed and let their mind freely wander (90 sec) and ending with a beep sound that instructed the subject to answer a short question about the content of their thoughts (Eryilmaz *et al.*, 2011).[1] The fMRI data was realigned using SPM8 and parcellated into $N = 90$ regions corresponding to the Automated Anatomic Labeling atlas (Fig. 5.9; Tzourio-Mazoyer *et al.*, 2002). Two regions were excluded because of signal dropout, leaving 88 regions. We estimated regional mean activity by averaging voxel time series in each brain region, and averaged across all subjects. We

[1] Data was acquired on a Siemens 3T Trio TIM in two scanning runs, TR/TE/FA = 1.1 s/27 ms/90°, matrix = 64×64, voxel size = 3.75×3.75×4.2 mm^3, 21 contiguous transverse slices, 1.05 mm gap, 2598 volumes.

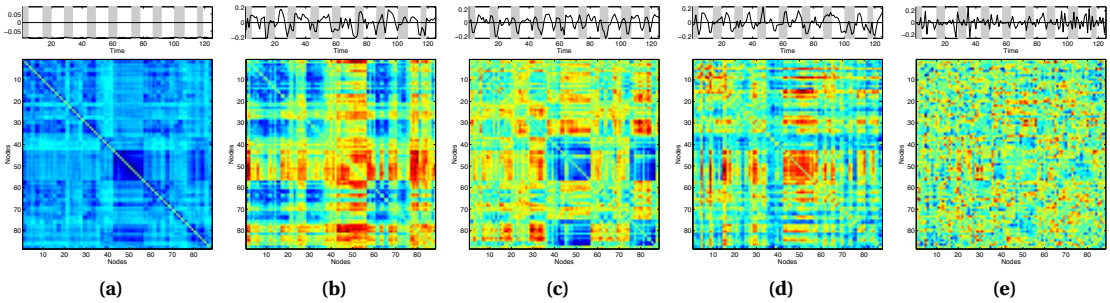


Figure 5.10 – Eigennetworks $\mathbf{S'}_{::k}$ for $k = 1, ..., 4, 100$. The top plots show the corresponding column $\mathbf{v}_{:k}$ overlaid on the stimulus paradigm (gray-shaded areas indicate the movie condition). The eigennetworks show the edge weights between 88 brain regions, ordered by lobe (frontal, limbic, occipital, parietal, subcortical, temporal) and with homologous regions adjacent to each other (e.g., left precentral cortex followed by right precentral cortex). The colorbars are symmetric around zero, i.e. red colors indicate positive edge weights, blue colors negative ones.

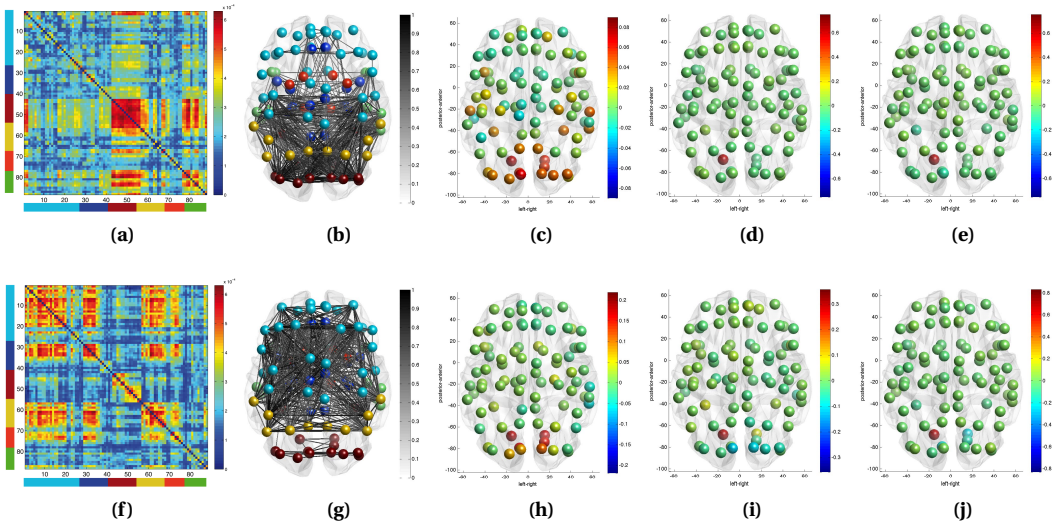


Figure 5.11 – (a) Adjacency matrix of the 88×88 brain graph built from the first and second eigennetwork, $\mathbf{A}'_1 = -0.1\mathbf{S'}_{::1} + 0.2\mathbf{S'}_{::2}$; (b) visualization of the graph in brain space, with nodes colored according to their anatomical location (light blue = frontal lobe, dark blue = limbic, red = occipital, yellow = parietal, orange = subcortical, green = temporal); (c-e) the energy of scaling and wavelet coefficients from coarse to fine scale, $J = 2$ scales and a Simoncelli-like design; (f) adjacency matrix of the graph built from $\mathbf{A}'_2 = -0.1\mathbf{S'}_{::1} - 0.3\mathbf{S'}_{::2}$; (g-j) as (b-e) and $-0.1 = 1/\sum_i \mathbf{v}_{i1}$.

constructed dynamic whole-brain functional networks by estimating temporal correlations between all pairs of brain regions using a sliding window approach with a window length of 30 scans. The absolute values of the correlation measures for each pair of regions were used as edge weights to construct $K = 125$ weighted graphs.

In this example, mode 3 captures the variability of the edges across time and the eigennetworks hence contain information on network dynamics. Fig. 5.10 shows eigennetworks $\mathbf{S'}_{::k}$ for $k = 1, .., 4, 100$ and their associated weights $\mathbf{v}_{:k}$. We observe well-known functional clusters in the average brain network $\mathbf{S'}_{::1}$ (Fig. a), such as the interconnected visual and auditory processing systems (nodes $43-56$ and $79-86$), the frontal attention network (nodes $7, 8, 29-34$), somato-motor cortex network (nodes

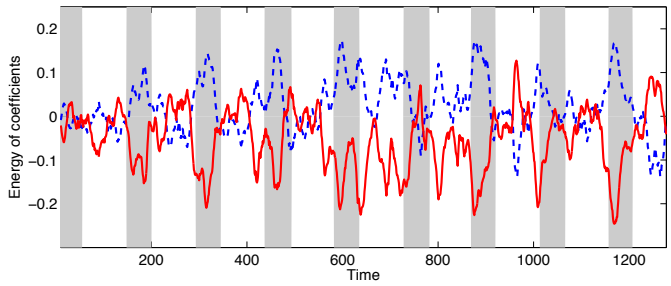


Figure 5.12 – Difference of the energy of the wavelet coefficients of the fMRI data decomposed with the SGWTs built from $\mathbf{A}'_1$ (resting condition) and $\mathbf{A}'_2$ (transitions of movie condition), $J = 2$ scales and a Simoncelli-like frame design. Difference of the energy for the scaling (red), and finest scale (blue dashed), where a positive value indicates a higher energy in the "resting" SGWT. A moving average with equal window length as for calculating the dynamic connectivity (i.e., 30) was applied, and gray-shaded areas indicate the movie condition.

1,2,19,20,57,58) and language-related areas (nodes $11-14$).

The other eigennetworks represent directions of variation from the average network across time. The second and third eigennetworks capture variability associated with the experimental paradigm, as can be seen form comparing $\mathbf{v}_{:2}$ and $\mathbf{v}_{:3}$ to the paradigm (Figs. b and c). Since the sign of $\mathbf{v}_{:2}$ is positive during the resting condition, edges with red color have larger weights during the resting condition, and those with blue color have larger weights during the transitions of the movies condition. For example, connectivity in the visual cortex is stronger during the resting condition, as previously observed (Richiardi *et al.*, 2011), and in frontal regions during the transitions of the movies condition.

As is the case when applying PCA or independent component analysis to noisy data, some of these components will represent noise (e.g., Fig. e depicting component 100). Discarding some of the eigennetworks could thus also help to reduce noise in a multislice graph.

To explore the multislice SGWT, we added and subtracted $\mathbf{S}'_{::2}$ from $\mathbf{S}'_{::1}$ to obtain the adjacency matrices $\mathbf{A}'_1$ and $\mathbf{A}'_2$, respectively (Figs. a,(b),(f),(g)). From $\mathbf{v}_{:2}$ we know that $\mathbf{A}'_1$ enhances edges typical for the resting condition, and $\mathbf{A}'_2$ for the transitions of the movies condition. $\mathbf{A}'_1$ has strongly connected visual and auditory cortices and a wavelet localized in the primary visual cortex diffuses to visual and auditory regions at coarser scales, as shown in Figs. c-(e). In contrast, as expected from the reduced connectivity within the visual cortex in $\mathbf{A}'_2$, the wavelet diffuses only within the primary visual cortex at coarser scales (nodes $43-48$), but not to extrastriate areas (nodes $49-56$), as shown in Figs. h-(j).

Instead of adding and subtracting $\mathbf{S}'_{::2}$ from $\mathbf{S}'_{::1}$, we could also have explored the strategy proposed in Eq. (5.26) with $\mathbf{g}$ equal to the stimulus paradigm to find the combination of eigennetwork weights that best approximates the stimulus paradigm.

We then implemented two SGWTs built from $\mathbf{A}'_1$ (resting) and $\mathbf{A}'_2$ (movies), respectively, and applied them to the normalized regional brain activity. We calculated the total energy of the wavelet coefficients at each scale and for each decomposition by summing over all brain regions. Due to energy conservation of the tight frame design, the ℓ^2-norm of the wavelet coefficients across all scales sums up to 1.

Fig. 5.12 shows the difference of the energy of the scaling (red line) and wavelet coefficients (blue dashed line), where a positive value indicates that the energy is higher for the resting than for the movies transform. We observe that at the finest scale, the energy is increased during the condition not well represented by the graph, i.e. the energy of the resting transform is larger than the movies

transform during the movies conditions (the blue line goes up). In previous work we estimated two separate brain networks for each condition (though not in a data-driven manner) and also observed an increase of energy when the SGWT was not adapted to the data (Leonardi & Van De Ville, 2011). For the scaling coefficients the picture is reversed, i.e. the energy of the resting transform is smaller than the movies transform during the movies condition (red line goes down).

Our results suggest that the variability identified by the HOSVD and illustrated with the eigennetworks, corresponds to relevant changes in network architecture across cognitive states. The column vectors of $\mathbf{V}$ give additional information regarding the importance of these networks across time, and the proposed multislice SGWT provides a meaningful decomposition of fMRI data.

5.1.6 Conclusion

We have extended the previously proposed SGWT to a tight frame, shown the link between classical and spectral graph wavelets, and presented a novel framework for the design of wavelets on multislice graphs based on the HOSVD. The proposed approach proceeds in two steps, by first estimating eigennetworks from an adjacency tensor and then building adjacency matrices that exploit the variability across slices of the network. Only then are these adjacency matrices converted to Laplacian matrices to apply the SGWT. The proposed framework of estimating eigennetworks and their combination is therefore not limited to the SGWT, but can be applied to other graph wavelet transforms (Coifman & Maggioni, 2006) or discrete signal processing frameworks (Sandryhaila & Moura, 2013) that are based on the adjacency matrix. Our results suggest that eigennetworks obtained from the HOSVD of a collection of adjacency matrices capture relevant variability across graph edges and that they provide meaningful graph wavelet transforms. In contrast to the common method of averaging or summing across multiple slices of a graph, the proposed framework allows to exploit the variability across them.

5.1.7 Appendix

Wave equation on graphs

The relation $\omega_\ell = \arccos(1-\lambda_\ell)$ between conventional Fourier frequencies and eigenvalues of the normalized Laplacian $\widetilde{\mathbf{L}}$ (i.e., $d_{\max}=1$) has also been found when considering the wave equation on complete and binary graphs (Von Below, 1985; Cattaneo, 1997; Friedman & Tillich, 2004):

$$\frac{\partial^2 u}{\partial t^2} = -\Delta u, \tag{5.27}$$

where Δ is a so-called edge-based Laplacian that acts as an operator on functions on the edges (in contrast to the conventional graph Laplacian that acts on signals defined on the nodes). The edge-based wave equation has a physical interpretation, where the edges are taut strings that are fused together at the nodes. While the node-based wave equation has infinite wave propagation speed, the edge-based wave equation has unit propagation speed (Friedman & Tillich, 2004).

From a purely signal-processing point-of-view, we can consider the wave equation Eq. (5.27) with the normalized graph Laplacian and discretized time. After applying the discrete temporal Fourier transform to Eq. (5.27), we retrieve the left-hand side as $(\cos(\omega)-1)u$. The equivalence with the eigendecomposition $\Delta u = \lambda u$ of the right-hand side then leads to $1-\cos(\omega)=\lambda$ with all eigenvalues contained in $[0,2]$.

5.2 HOSVD for multi-subject task-based dFC

In the journal article presented in this chapter we averaged the regional brain activity across all subjects. In Leonardi & Van De Ville (2013b), published in the proceedings of the IEEE International Workshop on Pattern Recognition in Neuroimaging (PRNI), we avoided the averaging step by adding a subject dimension to the adjacency tensor. The addition of a subject mode, as opposed to the temporal concatenation proposed in Chapter 3 is advantageous for task data because it exploits the temporal similarity across subjects, similar to tensor ICA (Beckmann & Smith, 2004). In this work we did, however, not exploit the graph wavelet transform. Instead, we studied how the eigennetworks were related to the experimental paradigm by regression of the associated time courses. Additionally, using leave-one-subject-out cross-validation, we investigated whether the eigennetworks generalized to unseen subjects and whether the experimental paradigm could be predicted from the strength of their expression.

5.2.1 Methods

Because the HOSVD is calculated from SVDs of matrix unfoldings of the tensor and modes 1 and 2 contain the same information, we can reduce the fourth-order *region* × *region* × *time* × *subjects* tensor to a third-order *connections* × *time* × *subjects* tensor. We derive this result for the third-order *region* × *region* × *time* tensor $\mathcal{A} = \mathcal{S} \times_1 \mathbf{U} \times_2 \mathbf{U} \times_3 \mathbf{V}$ studied in the journal article and the reduction of the fourth-order tensor follows analogously. First, we recall that the mode-3 matrix $\mathbf{V}$ is computed as the left singular matrix of the mode-3 matrix unfolding of $\mathcal{A}$: $\mathbf{A}_{(3)} = \mathbf{V}\Sigma\mathbf{W}^T$, where $\mathbf{A}_{(3)} \in \mathbb{R}^{K \times N^2}$, i.e. each row is a vectorized adjacency matrix. Next, we rewrite Eq. (5.20) with the eigennetwork core tensor $\mathcal{S}'$ as $\mathcal{A} = \mathcal{S}' \times_3 \mathbf{V}$ and unfold also this equation along mode 3: $\mathbf{A}_{(3)} = \mathbf{V}\mathbf{S}'_{(3)}$ (De Lathauwer *et al.*, 2000a). Finally, equating both matrix equations yields $\mathbf{A}_{(3)} = \mathbf{V}\Sigma\mathbf{W}^T = \mathbf{V}\mathbf{S}_{(3)}$, which shows that the rows of $\Sigma\mathbf{W}^T$ are the vectorized eigennetworks. Thus, the third-order tensor can be reduced to a *time* × *connections* matrix without any loss of generality and, additionally, because of the symmetry of the adjacency matrices, it is sufficient to retain only a vectorized version of their upper triangular entries.

As in the journal article we then decompose the *connections* × *time* × *subjects* tensor $\mathcal{A}$ into connectivity, time and subject-dependent loadings and a core tensor using a HOSVD (Fig. 5.13a):

$$\mathcal{A} = \mathcal{S} \times_1 \mathbf{U}_{\text{conn}} \times_2 \mathbf{U}_{\text{time}} \times_3 \mathbf{U}_{\text{subjects}}, \tag{5.28}$$

where $\mathbf{U}_{\text{conn}} \in \mathbb{R}^{C \times C}$, $\mathbf{U}_{\text{time}} \in \mathbb{R}^{T \times T}$ and $\mathbf{U}_{\text{subjects}} \in \mathbb{R}^{S \times S}$ are the orthonormal mode matrices, C is the number of connectivity pairs, T the number of windows and S the number subjects. Similar to before, the multiplication of the core tensor $\mathcal{S}$ with $\mathbf{U}_{\text{conn}}$ transforms its columns into eigennetworks, and we define $\mathcal{S}' = \mathcal{S} \times_1 \mathbf{U}_{\text{conn}}$.

Next, we remove the subject mode matrix on the right hand side by multiplying both sides with $\mathbf{U}^T_{\text{subjects}}$. We use the associativity of the tensor-mode multiplication (De Lathauwer *et al.*, 2000a), i.e., $(\mathcal{S}' \times_2 \mathbf{U}_{\text{time}} \times_3 \mathbf{U}_{\text{subjects}}) \times_3 \mathbf{U}^T_{\text{subjects}} = \mathcal{S}' \times_2 \mathbf{U}_{\text{time}} \times_3 (\mathbf{U}_{\text{subjects}}\mathbf{U}^T_{\text{subjects}})$, and the orthogonality of the mode matrices to obtain

$$\mathcal{A} \times_3 \mathbf{U}^T_{\text{subjects}} = \mathcal{S}' \times_2 \mathbf{U}_{\text{time}}, \tag{5.29}$$

The left-hand side now contains weighted dFC matrices and the right-hand side eigennetworks with

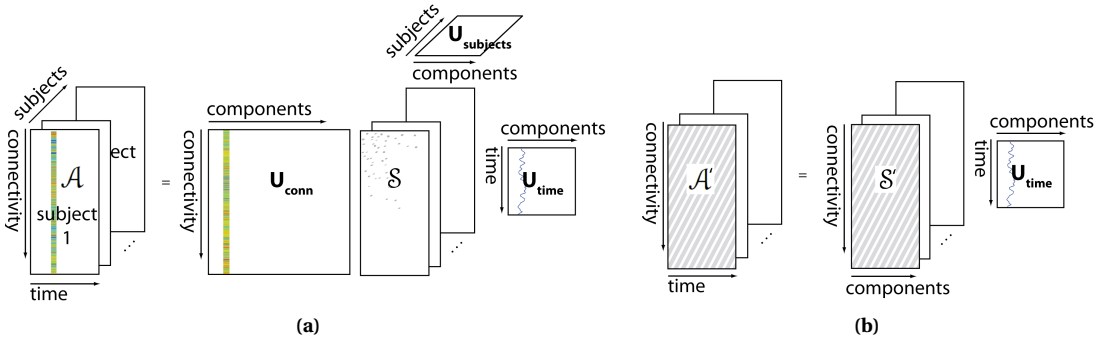


Figure 5.13 – (a) Multi-subject dFC modeled as a third-order *connections* × *time* × *subjects* tensor. HOSVD decomposes the tensor into a core tensor and connectivity, time and subject loadings. (b) The first frontal slice of $\mathcal{A}' = \mathcal{A} \times_3 \mathbf{U}_{\text{subjects}}^T$ and $\mathcal{S}' = \mathcal{S} \times_1 \mathbf{U}_{\text{conn}}$ contains average dFC and group eigennetworks, respectively.

associated time courses. We are interested in the *first* frontal slice of $\mathcal{A}' = \mathcal{A} \times_3 \mathbf{U}_{\text{subjects}}^T$ and $\mathcal{S}'$ because they contain dFC *averaged* across all subjects and *group* eigennetworks, respectively (Fig. 5.13b):

$$\mathbf{A}'_{::1} = \mathbf{S}'_{::1}\mathbf{U}_{\text{time}}^T. \tag{5.30}$$

We may write Eq. (5.29) in such a "slice-wise" notation using simple linear algebra. First, the matrix unfolding of Eq. (5.29) is $\mathbf{A}'_{(2)} = \mathbf{U}_{\text{time}}\mathbf{S}'_{(2)}$, where $\mathbf{A}'_{(2)} = [(\mathbf{A}'_{::1})^T|\ldots|(\mathbf{A}'_{::S})^T]$, and $\mathbf{S}'_{(2)} = [(\mathbf{S}'_{::1})^T|\ldots|(\mathbf{S}'_{::S})^T]$ (De Lathauwer *et al.*, 2000a). Second, this equation subsumes $(\mathbf{A}'_{::1})^T = \mathbf{U}_{\text{time}}(\mathbf{S}'_{::1})^T$. Thus, $\mathbf{A}'_{::1} = \mathbf{S}'_{::1}\mathbf{U}_{\text{time}}^T$.

The left-hand side of Eq. (5.30) contains *average* dFC: $\mathbf{A}'_{::1} = \sum_{s=1}^{S} u_{s1}^{\text{subjects}}\mathbf{A}_{::s}$ is a weighted average across all subject's dFC matrices $\mathbf{A}_{::s}$ and all $u_{s1}^{\text{subjects}} > 0$ because it is the first column of $\mathbf{U}_{\text{subjects}}$. In contrast to simple averaging, the weights u_{s1}^{subjects} are learnt from the data. Because $\mathbf{S}'_{::1}\mathbf{U}_{\text{time}}^T$ equals $\mathbf{A}'_{::1}$, we interpret the columns of $\mathbf{S}'_{::1}$ as *group* eigennetworks. Here, because of the additional dimension, only the time courses but not the eigennetworks are orthogonal.

To determine which eigennetworks were related to the task we regressed their time courses:

$$\mathbf{g} = \mathbf{U}_{\text{time}}\boldsymbol{\alpha} + \boldsymbol{\epsilon}, \tag{5.31}$$

where $\mathbf{g} \in \mathbb{R}^T$ was a demeaned moving average of the experimental paradigm (0 = rest, 1 = movies), $\boldsymbol{\alpha} \in \mathbb{R}^T$ the coefficients of interest and $\boldsymbol{\epsilon}$ a noise term (similar to the idea proposed in Eq. (5.26)). The solution to the ordinary least-squares problem is $\boldsymbol{\alpha} = \mathbf{U}_{\text{time}}^T\mathbf{g}$.

Finally, we wanted to examine how well the group eigennetworks and their combination as determined by Eq. (5.31) could predict the paradigm for an unseen subject. To answer this question, we applied the HOSVD within a leave-one-subject-out cross-validation loop. For all subjects $s = 1, \ldots, S$, we repeated the following three steps:

1. Solve Eq. (5.29) and Eq. (5.31), excluding dFC from subject s.
2. Project dFC of unseen subject onto the group eigennetworks estimated in step 1: $\mathbf{X} = (\mathbf{S}'_{::1})^{-1}\mathbf{A}_{::s}$. This equation solves for the time courses of the eigennetworks. Note that this problem corresponds to the inverse of Eq. (5.30): $\mathbf{U}_{\text{time}}^T = (\mathbf{S}'_{::1})^{-1}\mathbf{A}'_{::1}$.
3. Predict the experimental paradigm $\tilde{\mathbf{g}} = \mathbf{X}^T\boldsymbol{\alpha}$ for subject s, where $\boldsymbol{\alpha}$ are the coefficients learnt from all other subjects in step 1. We incrementally increased the number of included time courses, i.e.

the first prediction used only the time course of the first eigennetwork.

We compared the eigennetwork-based prediction against a FC-based one. In the competing approach we directly estimated the coefficients $\boldsymbol{\alpha}$ from the dFC time courses (averaged across mode 3 of $\mathcal{A}$ after excluding dFC from subject s), and $\mathbf{X}$ was simply equal to $\mathbf{A}_{::s}$. Here we incrementally increased the number of included pair-wise dFC time courses after ranking them by the absolute value of their coefficients $\boldsymbol{\alpha}$.

5.2.2 Results and discussion

Figs. 5.14a-d show the first few eigennetworks and their associated time courses. The time courses of the second and third eigennetworks strongly resemble the paradigm and more so than in Fig. 5.10, highlighting the advantage of including an additional subject dimension in the data structure.

The phase-randomization of dFC destroyed temporal similarity across subjects and the HOSVD estimated random eigennetworks and sinusoidal time courses (Figs. 5.14e-h). To investigate how these sinusoidal time courses arose, we generated three further types of surrogate datasets: 1) dFC estimated from phase-randomized regional time courses (preserving mean, variance and auto-regressive properties) yielded sinusoidal $\mathbf{U}_{\text{time}}$, but 2) dFC estimated from random, temporally uncorrelated noise (preserving mean and variance only) also resulted in sinusoidal $\mathbf{U}_{\text{time}}$. 3) adding random, temporally uncorrelated noise to mean dFC (preserving mean and variance of dFC but not its auto-regressive properties) yielded $\mathbf{U}_{\text{time}}$ with no readily visible structure. These results suggest that the sinusoidal fluctuations are a result of the autocorrelation in dFC that originates by the dFC estimation process itself (i.e., the sliding window operation).

The time-courses of the eigennetworks estimated from real dFC are not sinusoidal, however, because of the similarity of dFC across both time and subjects. Together, these results suggest that the HOSVD can efficiently and accurately extract changes in functional network configurations that are related to the experimental condition.

The group of connections identified in eigennetwork 2 is largely consistent with those shown in Fig. 5.10b. Notably, the red connections of the second eigennetwork, which are stronger during resting periods, are centered on higher visual regions (e.g. superior occipital cortex and fusiform gyrus with cuneus and lingual gyri). This indicates that while *activity* in visual regions was increased during movies as compared to the resting periods (Eryilmaz *et al.*, 2011), *FC* between visual regions was decreased during movies. The reduced FC of visual regions during movies is consistent with results of Richiardi *et al.* (2011) who studied static FC after concatenating all temporal segments corresponding to the same condition (Richiardi *et al.*, 2011). No knowledge of the experimental condition is required for the tensor representation of dFC and its decomposition. The blue connections, stronger during movies, were centered on the thalamus and frontal regions. Thus, for the thalamus, FC and activity were both increased during movies compared to resting periods (Eryilmaz *et al.*, 2011).

Interestingly, eigennetwork 4 showed the most striking similarity to the difference between average dFC during the two conditions (Fig. 5.14d and Fig. 5.14j). Its associated time course, however, suggests that this FC pattern may be driven by outliers and it did not contribute to out-of-subject predictions of the experimental paradigm (see also below).

The group eigennetworks generalized well to unseen subjects, as we could predict the paradigm for novel subjects after learning eigennetworks and regression coefficients from all other subjects (Fig. 5.14l). The folds-averaged correlation ($\pm$ std) between the predicted time course and the paradigm

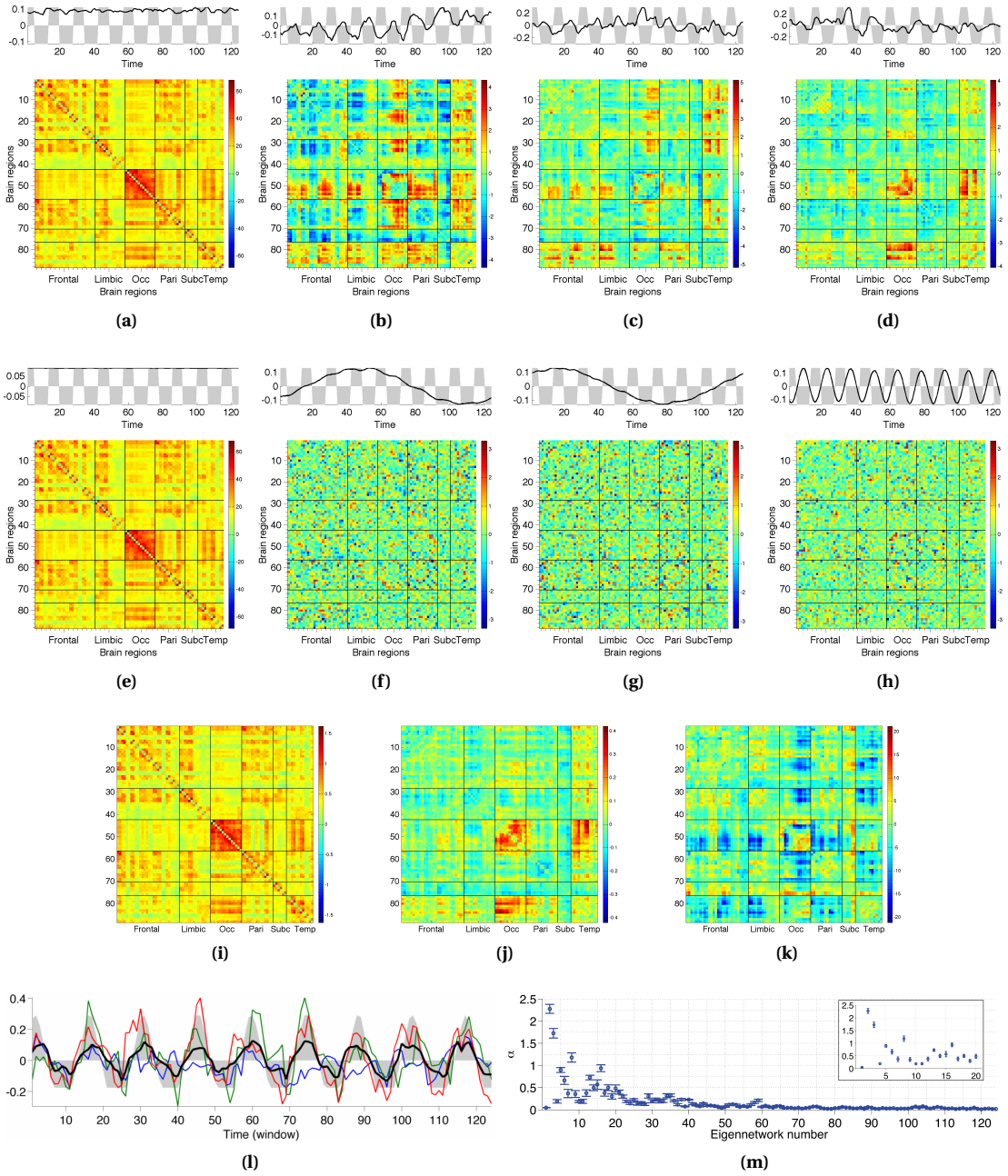


Figure 5.14 – (a-d) group eigennetworks $\mathbf{S'}_{:i1}$ with $i = 1,2,3,4,8$ and associated time courses (black) overlaid on a moving average of the experimental paradigm (gray shading); (e-h) the first four eigennetworks estimated from phase-randomized data; (i) average dFC across all windows and subjects, (j) difference between average dFC during movies and resting state. (k) Average linear combination of first 8 eigennetworks using weights $\boldsymbol{\alpha}$; (l) Predicted experimental paradigm using 8 eigennetworks for two “good" subjects (green and red lines) and one “bad" subject (blue line), and average prediction across all subjects (black line), overlaid on the true experimental paradigm (gray); (m) Coefficients $\boldsymbol{\alpha}$ averaged across all folds (absolute value because of sign invariance of HOSVD), error bars indicate standard error across folds and inset shows a zoom on the coefficients of the first 20 eigennetworks.

flattened at $\rho = 0.53 \pm 0.13$ for 8 components. This was mostly a result of including the second eigennetwork ($\rho = 0 \pm 0.22$ with the time course of the first eigennetwork and $\rho = 0.42 \pm 0.12$ with the first two). The second eigennetwork also had the largest weight across all folds, followed by eigennetworks 3 and 8 (Fig. 5.14m). The prediction was significantly worse when regressing the paradigm on all pair-wise connections, even for up to 100 included connections ($\rho = 0.18 \pm 0.20$). This demonstrates the benefit of dimensionality and noise reduction achieved by the HOSVD. The result that the HOSVD outperforms a prediction based on individual connections also indicates that a large-scale network, rather than a few individual connections, is reorganized during alternating movie and resting periods.

Finally, we estimated a difference network between movies and resting periods by calculating a weighted sum of the eigennetworks, with the weights $\boldsymbol{\alpha}$ determined from the regression of the time courses (Fig. 5.14k).The focus on FC of visual regions is now even more prominent. For example, FC between lower visual regions and, separately, between higher visual regions was high during movies (red connections), but FC between lower and higher visual regions was reduced, suggesting functional segregation during movies (blue connections).

5.3 Graph signal versus graph topology

In the journal article we showed a proof of concept that data-adapted and -unadapted graph wavelet transforms could succinctly reveal changes in cognitive states (Fig. 5.12). Co-activations of connected brain regions were most evident at coarse scales. At the finest scale, the energy of the wavelet coefficients was high if the graph was poorly adapted to the signal — i.e., the graph signal was picked up as isolated "high-frequency" noise. Based on this promising result, we explored whether graph wavelets could improve the scan-by-scan decoding of the two cognitive states by training classifiers, such as support vector machines, on graph wavelet coefficients from adapted and unadapted graphs (see Appendix D for more information). We, however, obtained similar prediction accuracies as when training classifiers on brain activity directly at the cost of a vastly more complex analysis pipeline (due to the estimation of functional networks and their wavelet transforms).

The underlying graph topology also influenced the signal representation in other ways than expected. The proposed analysis of BOLD co-activations was based on the idea that the graph wavelets act as zooms at different scales and locations. The underlying graph should therefore only serve the purpose of defining distance and thus what is considered close and far. However, for graphs with heterogeneous degree distributions, such as functional brain networks, both the eigenvalue spectrum and the coverage of the eigenvectors are unbalanced. To address the first issue, it worked well empirically to linearize the spectrum and Shuman *et al.* (2013) recently proposed a more elegant way to adapt the filter kernels to the distribution of the eigenvalue spectrum. This, however, destroys the direct mapping of eigenvalues to Fourier frequencies, but as the order of the eigenvectors remains unchanged, a conceptual link remains. The second issue led to an unequal coverage of brain activity. For example, activity of hubs tended to show up at fine scales and that of weakly connected brain regions at coarser scales. This made it difficult to separate the effects of graph topology and signal, and the link between the scale of analysis and the local/global extent of fMRI BOLD co-activation is obscured. This constituted a major drawback for the interpretation of the wavelet coefficients. For fMRI data, we therefore propose to focus the analysis either on the graph or on the graph signal.

In the former case, the eigenvectors of the Laplacian or the graph wavelets can directly be exploited

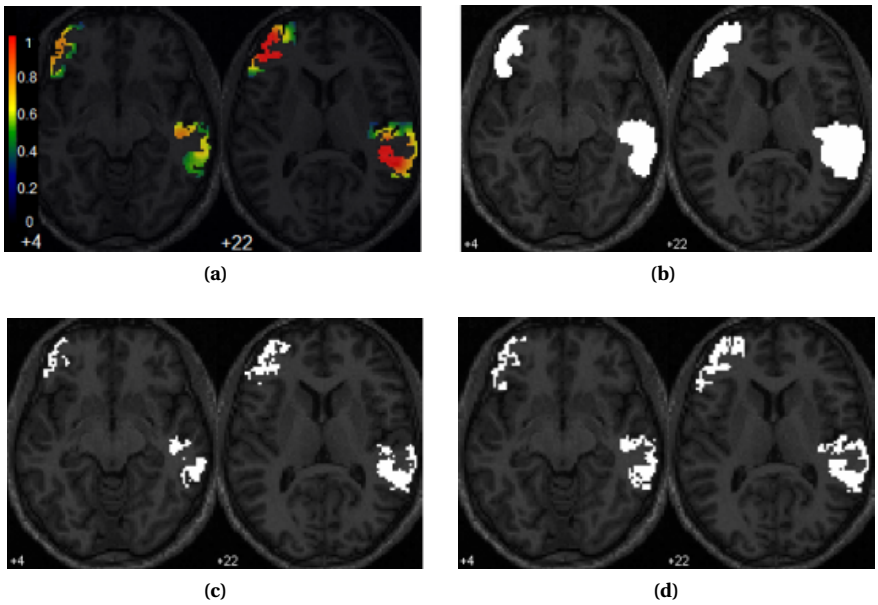


Figure 5.15 – (a) Simulated data before the addition of noise, (b) SPM contrast showing enhanced activity, (c) WSPM contrast, (d) graph WSPM contrast.

as features. Examples of this are the work presented in section 5.2, multi-scale community mining (Shen *et al.*, 2012; Tremblay & Borgnat, 2013), graph classification based on its first few Fourier coefficients (Hu *et al.*, 2013), and a diffusion model of the progression of Alzheimer's disease (Raj *et al.*, 2012).

In the latter case, we may define a fixed graph with a more regular geometry to obtain a more unbiased, multiscale view of the signal, while still benefitting from the ability of the graph wavelets to adapt to irregular geometries. This idea was exploited by Hamid Behjat, who completed his MSc thesis in our group and who extended wavelet-based statistical parametric mapping (WSPM) by replacing the classical wavelets with graph wavelets that dilate along gray matter and take boundaries into account (Van De Ville *et al.*, 2004; Behjat *et al.*, 2013). WSPM uses a spatial wavelet transform, which can efficiently represent clusters of activated voxels with only a few coefficients. Noise, on the other hand, will be evenly represented, so that a coefficient-wise t test can provide higher sensitivity than a voxel-wise test. Testing and thresholding in the wavelet domain also circumvents the need to heavily smooth the data as required in a traditional SPM analysis because of how multiple testing is corrected for. However, the wavelets WSPM is based on assume brain activity evolves on a regular Euclidean grid, hence the interest in incorporating better adapted wavelets.

For simulated data, graph WSPM detected many fewer false positives than SPM and its sensitivity was superior to that of conventional WSPM (Fig. 5.15). For single-subject auditory fMRI data, graph WSPM detected more finely structured activations in the auditory cortex than either of the other two approaches (Fig. 5.16). We also observed that adding coarser scales to the graph wavelet decomposition was more beneficial in terms of sensitivity than for conventional wavelets (Van De Ville *et al.*, 2007). Ongoing work by Hamid Behjat focuses on an extension to multi-subject studies. Another example of this idea is the recent work by Kim *et al.* (2014) who defined a cortical surface mesh and analyzed cortical thickness (the graph signal) to identify atrophied regions in Alzheimer's Disease patients.

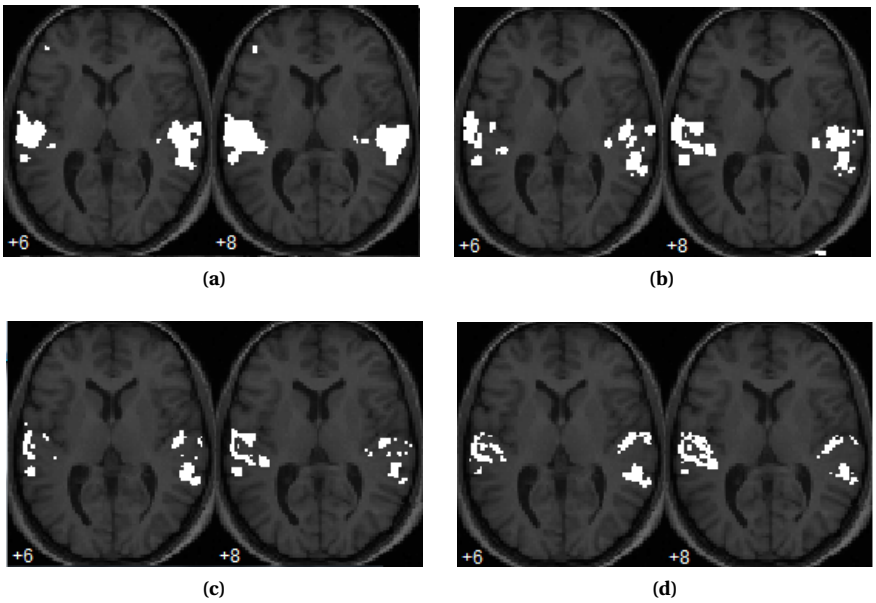


Figure 5.16 – (a) SPM contrast showing enhanced activity during auditory compared to resting periods, (b) WSPM contrast, (c) WSPM contrast with detections outside of GM boundaries removed, (d) graph WSPM contrast.

5.4 Outlook

In this chapter we introduced a tensor decomposition for the estimation of connectivity patterns from dFC and a graph wavelet transform for the multi-resolution analysis of brain activity. When there is no temporal similarity across subjects as during rest, the tensor decomposition can be simplified to an SVD, as explored in chapter 3. However, for task-based dFC, we can capitalize on this similarity to find network configurations that are consistent across subjects, while at the same time ignoring variability that is inconsistent and thus subject-specific.

Because the proposed tensor decomposition is invariant to any temporal reordering of the adjacency matrices, it implicitly assumes that all nodes are connected to themselves in all other layers. Because we were interested in recurring FC patterns, no matter when they occurred, the temporal ordering of the slices could be safely neglected in this work. As an alternative, if one wishes to keep this information, more complicated higher-dimensional tensors can be defined, which allow only certain inter-slice connections (De Domenico *et al.*, 2013). For example, for temporally ordered matrices, each node would only be connected to itself in the previous and following layers. As another alternative, "supra-adjacency" matrices have been proposed (Mucha *et al.*, 2010; Gómez *et al.*, 2013), where a new block-diagonal matrix of size $NK \times NK$ is built from all individual matrices. Off-diagonal entries can then represent links between layers. However, a graph with N nodes essentially becomes a graph with NK nodes and any results must be mapped back to the original N nodes. Kivelä *et al.* (2013) provide an excellent overview of multilayer networks and recent research.

Lastly, the proposed tensor decomposition identifies orthonormal mode matrices. This can be an unreasonable assumption for fMRI data, and in Chapter 4 we showed that the orthogonality constraint of PCA altered the topology of the FC patterns. For the construction of the graph wavelets the orthogonality constraint was beneficial, but if the estimated patterns are of main interest alternative

tensor decompositions provide interesting avenues for future work. Tensorial ICA is an obvious direction, while it would still impose statistical independence. Like tensorial ICA, CANDECOMP/PARAFAC computes a rank-1 decomposition but allows overlap between components. It has been successfully applied for facial recognition and the analysis of publication data (Vasilescu, 2011; Dunlavy *et al.*, 2011). For these alternative decompositions the strong spatial and temporal correlations of dFC are, however, potential issues. To reduce these correlations, the HOSVD could be applied in a first step, similar to the use of PCA before ICA in the matrix case.

6 Dynamic network reorganizations in healthy aging

In the previous chapters we estimated dFC with sliding-window correlations. While it is a simple and intuitive way to study dynamic brain networks, it has several inherent limitations: (i) One needs to choose a fixed window length, which limits the sensitivity to network changes. For example, with a window duration of 60 seconds changes on a faster time scale can not be resolved. (ii) Sliding-window correlations are inherently variable, that is it can fluctuate even for two signals with stationary or no correlation. (iii) We face a dimensionality problem: For a well-conditioned $N \times N$ correlation matrix, we need at least N time points. However, given usual TRs, choosing $T = N$ would drastically limit our ability to observe any changes in whole-brain dFC over time. For example, for $N = 80$ regions and a TR of 2 seconds, the window would cover almost 3 minutes. Thus, typically $T < N$ and the correlation matrix is rank-deficient.

Because the dFC estimates are noisy and have rank $< N$, it may be beneficial to constrain the patterns to simpler topologies. One way to simplify the patterns is to make them low-rank. The optimal low-rank approximation (in terms of explained variance) of a symmetric matrix $\mathbf{X}$ is computed from its eigenvalue decomposition $\mathbf{X} = \mathbf{V}\Lambda\mathbf{V}^T$, where $\mathbf{V} = [\mathbf{v}_1|\mathbf{v}_2|...|\mathbf{v}_N] \in \mathbb{R}^{N\times N}$ contains the eigenvectors and $\Lambda \in \mathbb{R}^{N\times N}$ is a diagonal matrix whose entries equal the associated eigenvalues $\lambda_1, \lambda_1, \ldots, \lambda_N$. The eigenvectors with the smallest eigenvalues are discarded and the matrix is approximated by the remaining L eigenvectors (rank-L approximation): $\mathbf{X} = \sum_{l=1}^{L} \lambda_l \mathbf{v}_l \otimes \mathbf{v}_l$, where λ_l is a scaling factor and $\otimes$ denotes the outer product (Fig. 6.1a). The most drastic simplification is a rank-1 approximation, that is $\mathbf{X} = \lambda_1 \mathbf{v}_1 \otimes \mathbf{v}_1$. Figs. 6.1b and c show real FC patterns and their rank-1 approximations. While not perfect, the approximations are surprisingly good. This visual impression is confirmed by high pair-wise spatial Pearson correlations ($r \in [0.70 - 0.87]$). The goodness-of-fit of low-rank approximations indicates that there is a lot of redundancy in the high-dimensional dFC space and resonates with results by Richiardi *et al.* (2010) who showed that low-rank approximations of static FC matrices are competitive with original FC matrices in terms of correctly predicting one of two cognitive states.

Importantly, in the case of rank-one FC patterns we no longer need to estimate $N(N-1)/2$ pair-wise correlations because each pattern is defined by only N values. This means we can directly identify patterns of brain activity from the *region* × *time* matrix and the window length is effectively reduced to 1 TR. We have essentially reverted back to the problem of identifying coherent fluctuations from spontaneous brain activity, which is typically solved using spatial ICA. However, because of

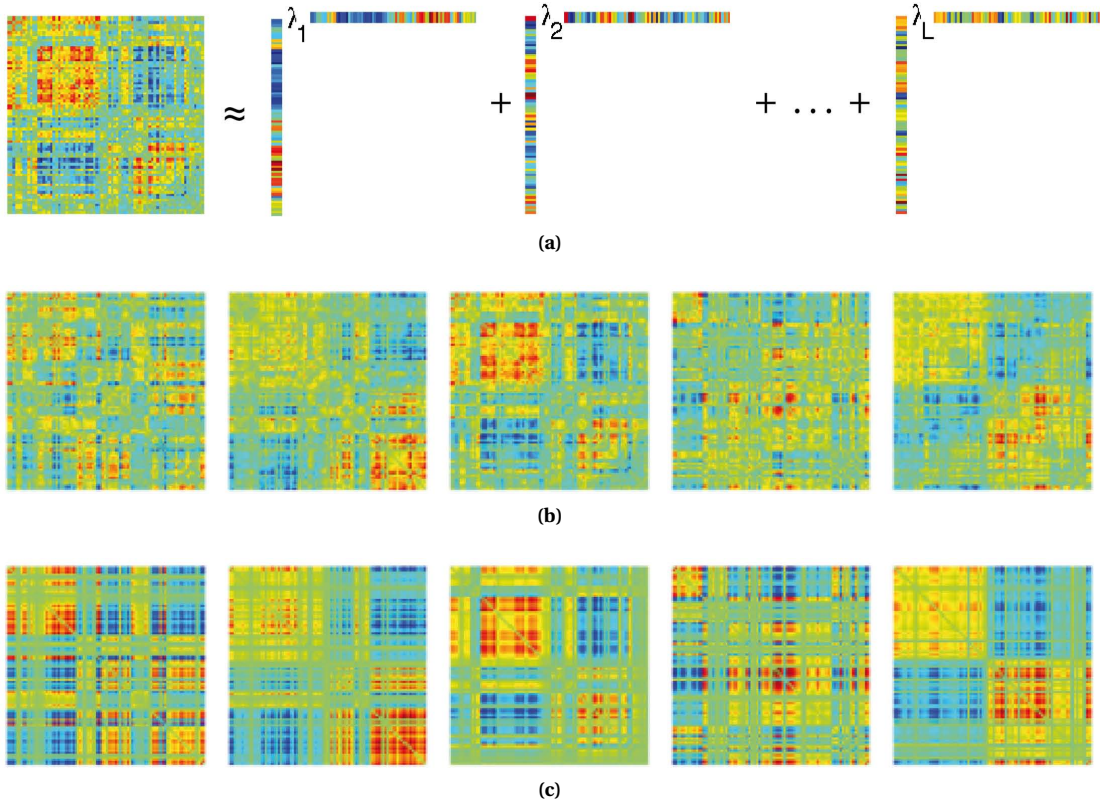
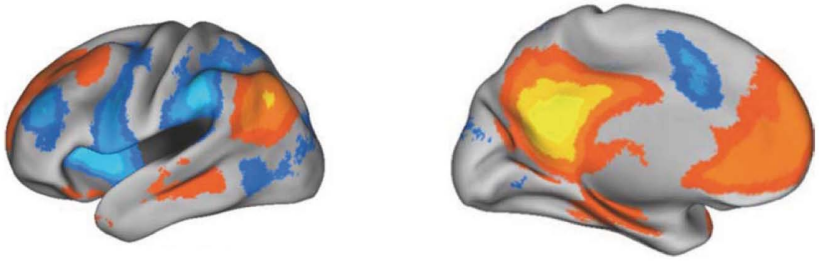


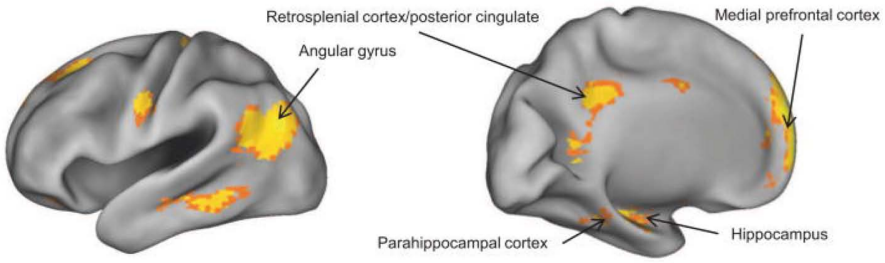
Figure 6.1 – (a) Approximating a matrix by a sum of L rank-one vectors; (b) FC patterns estimated from dFC of healthy, young subjects during rest, and (c) their rank-one approximations.

the constraint in space, the estimation of components is driven by spatial independence, which, as discussed earlier, provides an average view of RSNs. Instead, because a single brain region may participate in multiple networks over time, we desire a model that allows spatial overlap. Smith *et al.* (2012) reasoned that functionally distinct networks should also be largely temporally distinct and proposed temporal ICA to identify temporally independent, but potentially overlapping, RSNs. Overall, the identified independent components differed notably from traditional RSNs. An important consideration when applying temporal ICA is the TR of the data as the robustness of the algorithm is greatly increased by more time points, and the data used by Smith *et al.* (2012) had a short TR of 0.8 seconds. Also, spatial ICA was applied first to reduce the spatial dimensionality of the data to 142 regions. As both the authors and Calhoun *et al.* (2012) noted, independence in time may still be suboptimal for fMRI data and further relaxations of the time constraint may be needed.

An alternative proposal assumes temporal sparsity, or the exclusive occurrence of a single "co-activation pattern" (CAP) at any given scan (Liu & Duyn, 2013). To identify distinct CAPs Liu & Duyn (2013) temporally clustered the fMRI data into several groups using the k-means algorithm. Each CAP, or cluster centre, was then simply the average of all scans assigned to it. In k-means clustering, the clusters are estimated such that within-cluster distances are minimal, where within-cluster distances are defined as the distances of all data points to their cluster centre. Different distance functions can be employed but in this work the authors used 1 minus the Pearson correlation coefficient. Using the correlation will find scans with similar *patterns* of activity, while the Euclidean distance would be



(a)



(b)

Figure 6.2 – (a) The DMN (in red), reprinted from Buckner *et al.* (2008) with permission from John Wiley and Sons; and (b) regions activated during successful memory retrieval, reprinted from Rugg & Vilberg (2013) with permission from Elsevier.

influenced by *amplitude* differences in the BOLD signal over time. In the original proposal, based on prior work by Tagliazucchi *et al.* (2012), Liu & Duyn (2013) proposed to retain only those scans where the BOLD signal in a region of interest was above a threshold. This can be seen as a variant of seed correlation analysis. In later work they extended the analysis to include all fMRI scans (Liu *et al.*, 2013). Finally, we note that temporal clustering of fMRI data has already been proposed many years ago (e.g. Scarth *et al.*, 1995; Baumgartner *et al.*, 1997; Baune *et al.*, 1999) and growing interest in the temporal variability of patterns of brain activity has now brought these methods back to the foreground.

In this chapter we study transient CAPs in aging. Structural and average functional brain networks undergo prominent changes in both early and late stages of life. For example, Dosenbach *et al.* (2010) predicted the age of subjects aged 7 to 30 years from their FC and noted reduced strength of short-range FC between major RSNs with increasing age. At the other end of the age spectrum, Andrews-Hanna *et al.* (2007) reported reduced anterior-posterior FC in the DMN and Damoiseaux *et al.* (2008) reduced activity of two independent components, which they identified as anterior and posterior DMN. Also, both the efficiency (a measure of how densely connected the network is) and the modularity (the presence of strongly connected groups of regions) of functional networks are strongly affected by aging (Achard & Bullmore, 2007; Meunier *et al.*, 2009a). In accordance with the two studies cited above, frontal regions were more isolated in elderly subjects (Meunier *et al.*, 2009a).

Advanced aging is also associated with a decline in cognitive abilities, including memory loss. Because the DMN is thought to support episodic memory processing (Fig. 6.2; Greicius *et al.*, 2004; Buckner *et al.*, 2008; Spreng *et al.*, 2010; Rugg & Vilberg, 2013), the involvement of this network is of particular interest. While advanced aging has been consistently associated with a disruption of the DMN along an anterior-posterior axis (see also Ferreira & Busatto, 2013, for a recent review),

the relationship between FC and performance in neuropsychological tests (performed outside of the scanner) is less well understood. Several studies have reported associations of the integrity of the DMN with episodic memory (Andrews-Hanna *et al.*, 2007; Wang *et al.*, 2010; He *et al.*, 2012; Mevel *et al.*, 2013) and executive function (Damoiseaux *et al.*, 2008). All major regions of the DMN have been implicated, including the medial prefrontal cortex (mPFC), PCC, AG and hippocampus.

Additionally, slow-wave sleep is disrupted in older subjects (Van Cauter *et al.*, 2000). Slow-wave sleep is thought to be important for memory consolidation and stimulation of the PFC during sleep boosts memory performance (Marshall *et al.*, 2006; Diekelmann & Born, 2010). Recently, slow-wave activity in the mPFC during sleep, has been linked to both regional GM atrophy and memory retention (Mander *et al.*, 2013).

In this chapter we focus on co-activation patterns of the medial prefrontal cortex (mPFC) because of its important role in the DMN and its prominent structural and functional changes in advanced aging. We study the impact of normal aging on both average and transient patterns of coherent activity of the mPFC during rest and its influence on memory. In 24 younger and 33 older adults we first estimate average co-activation with the mPFC across the entire scan. While both groups showed robust co-activation of the default mode network (DMN), in the older subjects the mPFC significantly less often co-activated with posterior DMN regions, such as the medial temporal lobe, parahippocampus and retrosplenial cortex, in accordance with previous studies. Next, we studied transitory changes in mPFC CAPs in older subjects using temporal clustering. We observed 5 reproducible CAPs: the two most frequent CAPs strongly resembled the DMN, with differences in limbic regions, while the additional 3 CAPs showed co-activation of the mPFC with bilateral insulae, extensive frontal regions PCC and AG only, and visual and motor regions, respectively. In young subjects, the mPFC co-activated with other regions of the DMN most of the time. The occurrence rate of the frontal/insular CAP in elderly subjects negatively correlated with memory retention, and the association remained after correcting for age, head motion and gray matter volume in the mPFC. Our results highlight flexible, transitory interactions between the mPFC and large-scale brain systems and suggest their functional relevance in healthy aging.

6.1 Introduction

Cognitive functions, such as memory, attention and task-switching, decline in healthy aging. Neuroimaging studies have been crucial to better understand brain mechanisms that underlie such changes by identifying the location of structural and functional changes. Structural imaging has revealed strong gray matter (GM) atrophy, particularly in frontal and parietal areas, as well as reduced prefrontal white matter integrity (Good *et al.*, 2001; O'Sullivan *et al.*, 2001; Sowell *et al.*, 2003; Gunning-Dixon *et al.*, 2009; Giorgio *et al.*, 2010). Functional imaging has indicated that older adults engage frontal areas more strongly than younger adults over a range of tasks, but dis-engage a set of regions known as the "default mode network" (DMN) less strongly (Grady *et al.*, 1994; Lustig *et al.*, 2003; Spreng *et al.*, 2010). The DMN comprises a group of regions that show increased and coherent activity (functional connectivity, FC) while a subject is at rest (Raichle *et al.*, 2001; Greicius *et al.*, 2003), and is thought to be implicated in self-referential and episodic memory processing (Greicius *et al.*, 2004; Buckner *et al.*, 2008). The strength of coherent spontaneous fluctuations in the DMN gradually declines with age (Andrews-Hanna *et al.*, 2007; Damoiseaux *et al.*, 2008; Meunier *et al.*, 2009a; Mevel *et al.*, 2013). This disruption appears to preferentially occur along an anterior-posterior axis, involving the major hubs of the DMN: the posterior cingulate cortex (PCC) and medial prefrontal cortex (mPFC; Andrews-Hanna *et al.*, 2010). While it is less clear how age differences relate to cognitive performance, a few studies have reported positive relationships between the FC among DMN regions and memory (Andrews-Hanna *et al.*, 2007; Wang *et al.*, 2010; He *et al.*, 2012; Mevel *et al.*, 2013) and executive function (Damoiseaux *et al.*, 2008).

In addition, slow-wave sleep, which is important for memory consolidation, is reduced in older adults (Van Cauter *et al.*, 2000; Diekelmann & Born, 2010). Enhancing slow oscillations ($< 1Hz$) during sleep by electrically stimulating the PFC boosts memory performance in young adults (Marshall *et al.*, 2006) and reduced prefrontal slow-wave activity in older adults is related to impaired memory retention (Mander *et al.*, 2013).

Previous studies on resting-state FC implicitly assumed stable functional interactions across time. Functional interactions are, however, highly dynamic (Chang & Glover, 2010; Hutchison *et al.*, 2013b), and the estimation of a single network across time provides only an average view of these interactions. To reveal patterns of transient functional interactions several alternative techniques have recently been proposed (Kiviniemi *et al.*, 2011; Smith *et al.*, 2012; Liu & Duyn, 2013; Leonardi *et al.*, 2013; Eavani *et al.*, 2013; Allen *et al.*, 2014). These results suggest that many regions are part of multiple networks across time (at least as resolved by fMRI). The relationship of transient interactions with behavior are, however, even less well understood than for average interactions. A few studies have shown that the network topography just prior to a task is predictive of errors (Eichele *et al.*, 2008; Ekman *et al.*, 2012) and reaction time (Thompson *et al.*, 2013b). That transient network organizations during rest have functional relevance is suggested by recent work with clinical populations (see e.g. Hutchison *et al.*, 2013a, for an overview). For example, we observed changes of transient FC in parietal and prefrontal regions in patients with multiple sclerosis (Leonardi & Van De Ville, 2013b), and Jones *et al.* (2012) reported altered occurrences of DMN subnetwork configurations in Alzheimer's patients. Estimating only average interactions may thus ignore important temporal variability and differences in transient interactions may in fact explain changes observed in average FC.

The goal of this study was to determine whether changes in transient network configurations underlie age-related changes commonly observed in average networks. Given the prominent role of

mPFC in MRI and sleep studies on healthy aging and memory, we focus our analysis on its co-activation patterns (CAPs). Each CAP describes a unique spatial pattern of frequently co-activating regions and is identified by temporally clustering the fMRI data (Liu & Duyn, 2013). FMRI scans are assigned to the CAP they are most similar to, a simple and intuitive approach. Temporal clustering of fMRI data has already been proposed much earlier for task-based fMRI data (e.g. Scarth *et al.*, 1995; Baumgartner *et al.*, 1997; Baune *et al.*, 1999), but is especially relevant today in light of the evidence of temporal network variability. Because still little is known about the behavioral significance of changes in transient interactions, we also examined whether their occurrence was associated with individual differences in episodic memory in older adults.

6.2 Methods

Participants

Study participants included 24 young adults and 34 older adults recruited from the students and community of Stanford University. Written, informed consent was obtained from all participants and the Stanford University Institutional Review Board approved the study protocol. Data of the young adults was previously analyzed using traditional FC fMRI analyses (Shirer *et al.*, 2012). Memory retention in the older adults was assessed with percentage retention of the Hopkins Verbal Learning Test Revised (HVLT-R PR, Brandt, 1991; Benedict *et al.*, 1998). Subjects were asked to recall words from a list of 12 words, both immediately and after a 20-minute delay. HVLT-R PR is calculated as the ratio of delayed to immediate memory. Data from 1 older adult was excluded because of excessive motion during fMRI (mean displacement exceeded 1 mm). All participants had mean frame-wise displacement (FD; Power *et al.*, 2012) < 0.5 mm. Table 6.1 shows demographics of the 24 young and 33 older adults included in this study.

Data acquisition

A structural and resting-state scan were acquired during the scanning session. For the young participants, only 13 received a structural scan. Functional images were acquired on a 3.0T GE scanner (TR = 2s, TE = 30 ms, flip angle = 80°, 1 interleave, matrix size 64 x 64, FOV = 22 cm). Heart and respiration rates were monitored during the scan. A high-resolution structural scan was acquired using an axial 3D fast spoiled gradient recalled echo sequence (FSPGR; 162 slices, 0.86-mm2 in-plane and 1-mm through-plane resolution, TR = 5.9 ms, TE = 2 ms, flip angle = 15°, FOV = 22 cm).

Young adults were scanned for 10 minutes and older adults for 8 minutes while resting with their eyes closed. Data from young adults was shortened to 8 minutes for all analyses. A debriefing confirmed that all subjects stayed awake throughout the scan.

Structural MRI analysis

To control for the possible effect of GM atrophy on resting-state co-activations, we used gray matter (GM) volume in the seed region of interest (defined below) as a covariate in our analyses. We measured GM volume using FSL's voxel-based morphometry. After removal of non-brain structures, the high-resolution structural scans were segmented into GM, WM, and CSF. Because our study included more elderly than young participants (of which only 13 had a structural scan), we selected a subset of 13

	Young	Old
N	24	33
Percent male	37.5	39.4
Age (y)	23.5 ± 3.8	66.8 ± 8.4
Education (y)	n.a.	16.4 ± 1.8
MMSE	n.a.	29.8 ± 0.5
HVLT-R PR (%)	n.a.	93.1 ± 9.9
RMS (mm)	0.18 ± 0.10	0.16 ± 0.07
FD (mm)	0.015 ± 0.005	0.023 ± 0.016**
GM density mPFC (au)	0.70 ± 0.05‡	0.48 ± 0.07**

y, year; MMSE, Mini Mental State Examination; HVLT-R PR, Hopkins Verbal Learning Test Revised percent retention; RMS, root-mean square displacement; FD, frame-wise displacement; au, arbitrary unit; ** indicates a significant age difference at p<0.01
‡ Available for 13 young subjects

Table 6.1 – Participant information.

older adults to create a study-specific template. The demographics of this subset of subjects were similar to those of all subjects (5/8 m/f, mean ± SD age 65.7 ± 10.6 years, education 16.3 ± 1.8 years, MMSE 29.8 ± 0.4, compare to Table 1). The GM images of the 26 subjects were transformed into MNI152 standard space using a non-linear registration and a study-specific template was created. Native GM images of all subjects were then nonlinearly registered to the template, modulated and smoothed with a 7-mm Gaussian kernel. Finally, we averaged GM density in the region of interest (see below).

Preprocessing of resting state data

Data were preprocessed using the following steps with FMRIB's Software Library (FSL version 4.1): Removal of the first 6 scans to allow the MR signal to equilibrate; physiological noise reduction using RETROICOR and RVHRCOR (Glover *et al.*, 2000; Chang & Glover, 2009); rigid head motion correction; normalization to the MNI152 standard space at 3x3x3-mm resolution; spatial smoothing with a 6-mm FWHM Gaussian kernel; removal of the linear trend to correct for drift and regression of the 6 head motion parameters, mean cerebrospinal fluid (CSF) and white matter (WM) signals; band-pass filtering (0.01-0.1 Hz); and despiking of each voxel's time course by removing outliers greater than 3 standard deviations from the voxel's mean to further reduce the impact of head motion (similar to AFNI's 3dDespike). Finally, the signal of each voxel was z-scored so that the data could be compared across subjects.

Co-activation pattern analysis

We used a region of interest in the ACC/mPFC, centered at [-3 47 16] similar to the mPFC seed location in Andrews-Hanna *et al.* (2007) (Fig. 6.3). The seed region was defined by spatially clustering resting-state fMRI data of an independent dataset (Ng, unpublished).

The estimation of average and transient CAPs follows the steps described in Liu & Duyn (2013) and is shown schematically in Fig. 6.3. First, we extracted fMRI scans with supra-threshold signal ($z > 1$) in the mPFC seed region. Liu *et al.*'s idea to reduce the fMRI data to a few select scans, which retain enough spatiotemporal information to recapitulate (average) resting-state networks, was largely inspired by previous work of Tagliazucchi *et al.* (2012). As opposed to Liu & Duyn (2013) we also retained scans with

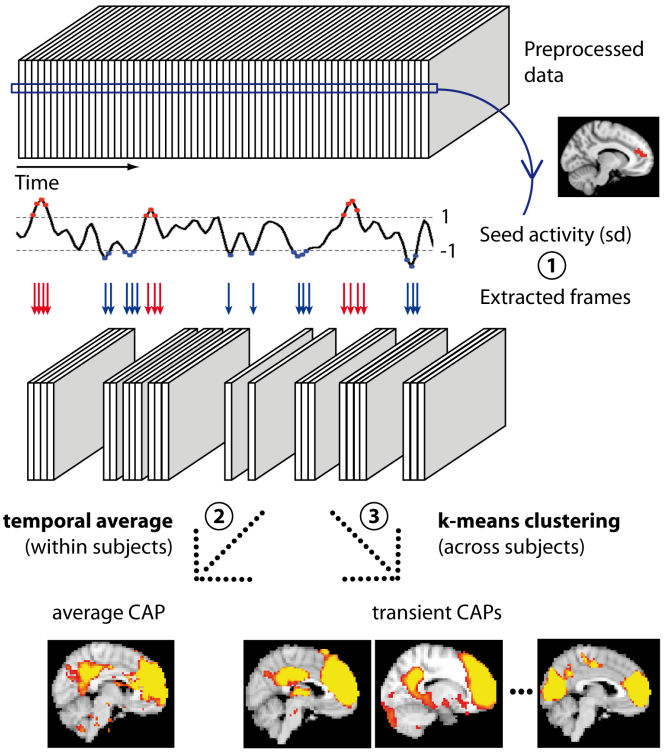


Figure 6.3 – Illustration of analysis pipeline: scans where BOLD activity of the mPFC/ACC seed region exceeds 1 standard deviation are extracted. The sign of all sub-threshold scans (z<-1) is flipped to convert deactivations to activations. Average and transitory mPFC-coactivation patterns are estimated by averaging across all extracted scans and temporally clustering them, respectively.

sub-threshold signal ($z < -1$) but flipped their signs to convert deactivations to activations (otherwise k-means clustering would assign these scans to separate clusters even if the pattern of co-activation is the same).

Second, to estimate average CAPs we simply averaged the extracted scans for each subject. We compared average CAPs of young and older subjects by applying voxel-wise nonparametric permutation testing (FSL's randomize using threshold-free cluster enhancement with p<0.05 family-wise error corrected). Because older subjects moved significantly more inside the scanner (see Table 1), we included mean FD as a nuisance regressor.

Third, in older subjects only, we concatenated the extracted scans across time and clustered them into several distinct co-activation patterns using k-means (correlation distance function). We only included elderly subjects in this analysis because of unequal group sizes and to specifically focus on transitory changes in mPFC coactivations in older age that may underlie changes in average CAPs. The k CAPs were obtained by averaging all scans assigned to the same cluster. The number of clusters was determined using reproducibility resamplings. Specifically, the 33 subjects were randomly split into two groups of 16 and 17 subjects each, and for each group the extracted scans were temporally concatenated and repeatedly clustered with $k = 2, 3, \ldots, 12$. Reproducibility was quantified by spatially correlating the CAPs from the two independent datasets. We performed 100 split-halves and for each split retained the Pearson r of the least reproducible CAP (worst "best match") as a measure of minimal reproducibility. We compared the reproducibility at each value of k to a surrogate dataset, which

was generated under the stationarity assumption (i.e., a single CAP occurs over time). The surrogate dataset was created by adding white noise with the same variance as the real data to a subject-specific stationary CAP, which was randomly scaled over time. Because each subject's fMRI data has zero temporal mean, the stationary CAP was estimated from the first eigenvector of its spatial covariance matrix. The outer product of the first eigenvector is a rank-one, and thus stationary, approximation of the covariance matrix. We estimated the reproducibility of k=2 (i.e., one too many) clusters for 1000 surrogate datasets. We retained CAPs whose average reproducibility was above the 95th percentile of the reproducibility in surrogate data.

The percentage occurrence of each CAP was obtained by dividing the number of scans assigned to each cluster by the total number of extracted scans, and the within-cluster similarity was obtained by estimating the spatial correlation of each individual scan with its CAP and then averaging the correlations for each CAP. Next, we quantified the fractional occurrence of each reproducible CAP in each subject by dividing the number of scans assigned to each cluster by the subject's total number of extracted scans. We also estimated occurrence of these CAPs in young subjects by assigning each extracted scan ($|z| > 1$) to the closest CAP estimated from the elderly subjects. Scans whose distance to the closest cluster was larger than the 95th percentile of within-cluster distances in elderly subjects were assigned to a separate cluster. We compared the occurrence of CAPs between young and old healthy subjects using 2-sample t tests. Significance was assessed with permutation-based nonparametric testing (1000 random assignments to two groups of 24 and 33 subjects, corrected for multiple comparisons by retaining the maximum t statistic). Additionally, to validate this assignment, we also temporally clustered the retained scans retained of the young subjects.

To assess whether there was a relation with performance on the memory task, we estimated the correlation between the HVLT-R PR score and the fractional occurrence of each CAP, controlling for age, mean FD and GM volume of the mPFC seed region.

6.3 Results

Correspondence of co-activation and seed correlation maps

Fig. 6.4 shows the mPFC seed correlation map (a) and average CAPs estimated from all scans where signal in the mPFC seed region was supra threshold (b), sub threshold (after flipping the sign, c), and either supra or sub threshold (d) for a representative young subject. The average in (b) was calculated from only 35/234 (15%) scans and in (c) from 27/234 (12%) scans. Not only the average of above threshold scans closely replicates the seed-correlation map, in accordance with Tagliazucchi *et al.* (2012) and Liu & Duyn (2013), but also the average of sub threshold scans. The spatial Pearson correlation coefficient r was 0.89 or higher for all pair-wise comparisons ($p < 10^{-10}$ corrected, Fig. 6.4e). Correspondence between all four maps was excellent across all subjects (Fig. 6.4f). While the supra and sub threshold CAPs were least similar, their correlation was still high ($r = 0.72 \pm 0.1$), and the combined supra and sub threshold scans almost perfectly replicated the seed correlation map ($r = 0.98 \pm 0.01$, all $p < 10^{-10}$ corrected).

On average across subjects, BOLD signal in the mPFC was supra or sub threshold in 11% of scans (Fig. 6.4g; mean ± SD, $z > 1, 11.4 \pm 1.7\%; 11.5 \pm 1.5\%; z < -1, 11.1 \pm 1.7\%, 12.2 \pm 1.5\%$, for young and old subjects, respectively). There were no significant differences between young and elderly subjects ($p > 0.05$). Using both supra and sub threshold scans, thus approximately doubled the number of

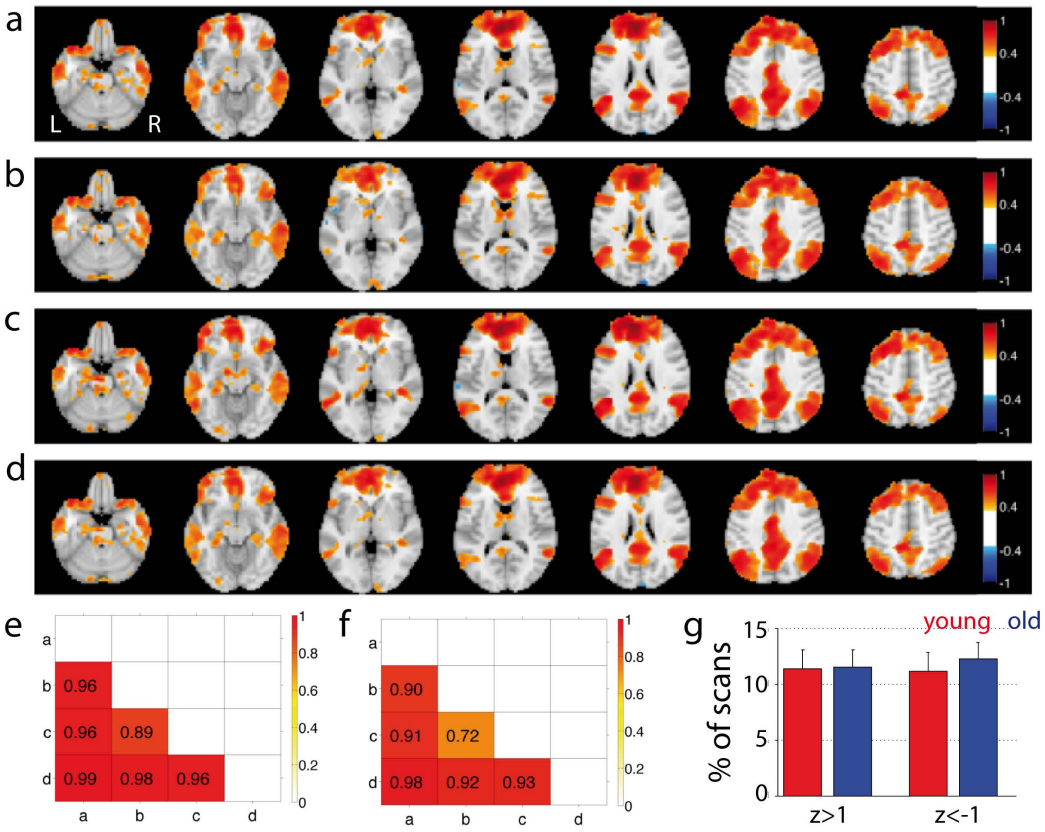


Figure 6.4 – (a) mPFC seed-correlation map. Average of (b) supra threshold scans ($z > 1$), (c) sub threshold scans ($z < -1$), and (d) combined supra and sub threshold scans ($|z| > 1$) for an example subject. The sign of sub threshold scans was flipped before all calculations. The color maps were adjusted by dividing by the maximum value for the purpose of comparison. Slices at z = [-26 -14 -2 10 22 34 46]. (e) Spatial correlation coefficient between the maps shown in a-d. (f) Average correlation coefficient across all subjects. (g) Percentage of scans with supra and sub threshold signal in the mPFC in young and old subjects, errorbars are standard deviations across subjects.

retained scans and improved the correspondence with seed correlation maps.

Average CAPs of young and elderly subjects

The average CAP in young subjects showed prominent co-activation of DMN regions: PCC, retrosplenial cortex (Rsp), bilateral angular gyri (AG), thalamus, medial temporal lobe (MTL), temporal pole, parahippocampal cortices (PHC) and hippocampal formation (HF) (Fig. 6.5a).

The average CAP in elderly subjects was similar overall (Fig. 6.5b), but showed significantly reduced co-activation of inferior parietal and temporal regions, notably the bilateral AG, medial temporal pole, PHC and HF (Fig. 6.5c).

Transient CAPs in elderly subjects

To reveal transitory mPFC co-activation patterns, which may remain hidden in average analyses, we temporally clustered all extracted scans of elderly subjects into several CAPs. Based on a comparison of the reproducibility of the CAPs to a surrogate dataset under the stationarity assumption, we identified

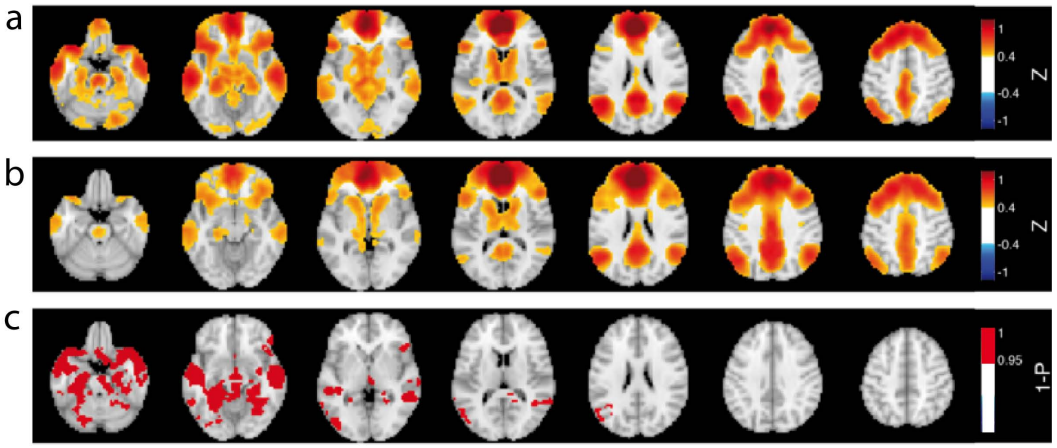


Figure 6.5 – Average mPFC-CAP in (a) young and (b) older subjects. (c) Significantly more co-activated voxels in younger subjects (FWE-corrected p-value <0.05). There were no significantly less co-activated voxels in younger subjects. Slices at z = [-26 -14 -2 10 22 34 46].

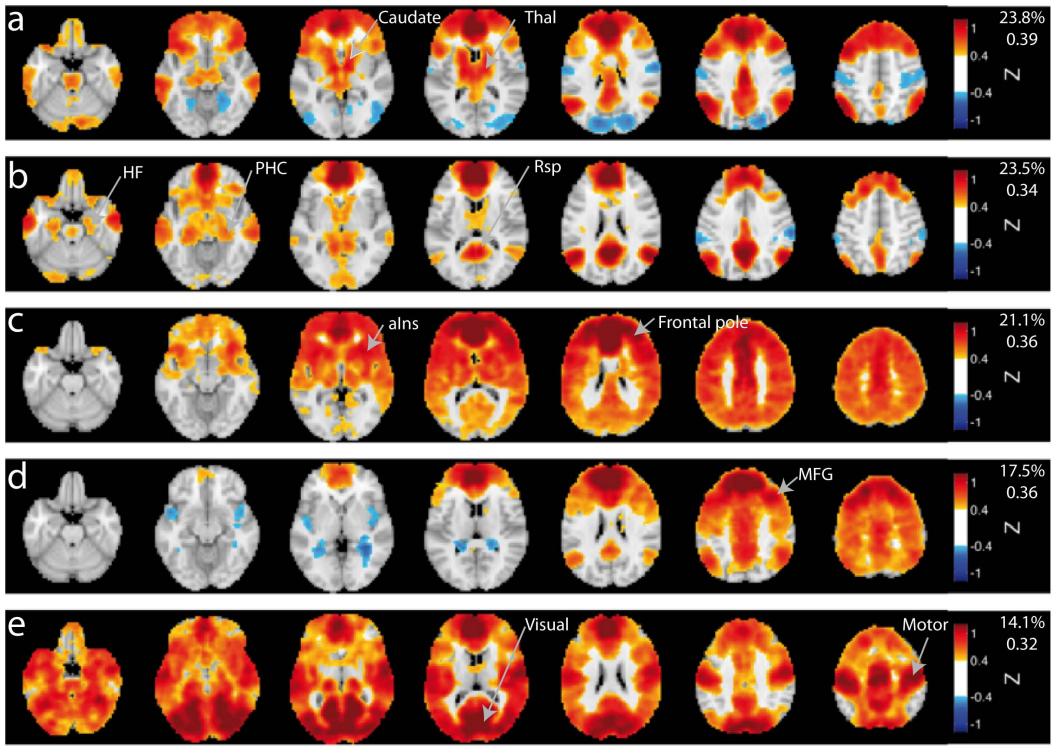


Figure 6.6 – (a-e) 5 reproducible CAPs were observed in older subjects: (a) and (b) strongly resemble the DMN, (c) includes orbitofrontal cortex and frontal pole, (d) superior regions of the DMN, and (e) primary motor and visual areas. The percentages of occurrences are indicated in the upper left corner. Slices at z = [-26 -14 -2 10 22 34 46], CAPs are ordered by percentage occurrence, indicated next to the colorbar. Within-cluster similarity is indicated below the percentage occurrence.

5 reproducible CAPs. The transient CAPs showed striking differences (Fig. 6.6, sorted by decreasing occurrence): The first two CAPs strongly resembled the DMN. Specifically, they both showed co-

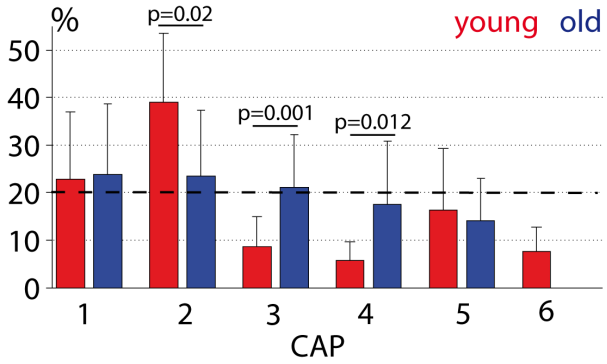


Figure 6.7 – Occurrence rate of the CAPs in elderly and young subjects. CAP 6 contains scans, which were too far from any CAP identified in old subjects. Errorbars indicate standard deviation, p values are corrected for multiple comparisons and after controlling for mean FD. The dashed line indicates equal occurrence of all CAPs (20%).

activation of posterior midline, inferior parietal, medial temporal and parahippocampal regions. In addition CAP 1 included the thalamus, caudate and orbitofrontal cortex (OFC); and CAP 2 the Rsp, HF and PHC. Both CAPs also showed weak deactivations covering the postcentral gyrus and occipital cortex. The remaining CAPs differed markedly. CAP 3 covered extensive frontal areas, including the middle (MFG) and superior frontal gyrus (SFG), frontal pole and anterior insula, and extending into the middle cingulate cortex. This CAP was largely devoid of temporal areas. CAP 4 included posterior midline and parietal but not temporal DMN regions, as well as the mPFC and MFG. The strongest contributions in CAP 5 were from visual and motor cortices.

Next, we estimated the occurrence of these CAPs in the young subjects by assigning their scans to the closest CAP (7.5% of the scans remained unassigned because they were too far from any CAP). Percentages of occurrence for each CAP are shown in Fig. 6.7. The mPFC co-activated with other DMN regions a majority of the time in young subjects (CAPs 1 and 2 together occurred 52% of the time) and the occurrence of CAP 2 was significantly increased in young subjects (mean ± SD occurrence $39.0 \pm 13.8\%, 23.5 \pm 14.8\%$ in young and old subjects, respectively, Cohen's $d = 0.73$, $p = 0.02$). In contrast, the occurrences of CAPs 3 and 4 were significantly reduced in young subjects (CAP 3: $8.9 \pm 6.8\%, 21.1 \pm 11.0\%$, Cohen's $d = -0.90$, $p = 0.001$; CAP 4: $5.9 \pm 4.0\%, 17.5 \pm 13.3\%$, Cohen's $d = -0.75$, $p = 0.001$). Additional results obtained by temporally clustering the extracted scans of the young subjects confirmed that the CAPs of young subjects strongly resembled the DMN, that the motor-visual CAP was preserved, and that the two frontal CAPs had poor correspondence (Fig. E.1). While CAP 4 in young subjects also showed co-activation of the insula and prefrontal areas, it additionally included posterior regions of the DMN.

Finally, we assessed whether there was a relationship between the occurrence rate of the different CAPs in individual subjects and memory. The percentage occurrence of CAP 3 was significantly negatively correlated with memory retention ($r = -0.50$, $p = 0.0027$, survives Bonferroni correction for multiple comparisons, Fig. 6.8). The number of extracted scans and the occurrence of the other 4 CAPs did not correlate with memory.

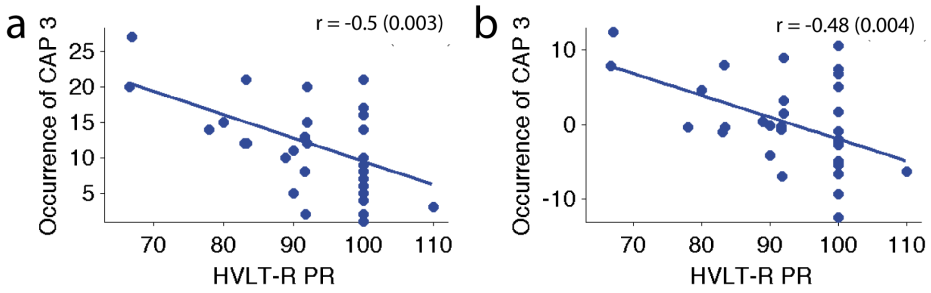


Figure 6.8 – (a) Percentage occurrence of CAP 3 plotted against memory and regression slopes for the correlation. (b) Same as (a) after controlling for age, mean FD and GM density in the seed. r is Pearson correlation coefficient and p value is indicated between brackets.

6.4 Discussion

Our study revealed that transient CAPs of the mPFC are linked to age and episodic memory. First, the occurrence of CAP 2 was decreased in older adults. This pattern resembled the DMN and included the HF and PHC, which are important in episodic memory processing. Second, the occurrence of two patterns, which were rarely observed in younger adults and notably devoid of temporal regions of the DMN, was increased. Third, the occurrence of a CAP with prominent contribution of prefrontal areas in individual subjects significantly predicted worse performance in a memory task. Our findings suggest a selective and detrimental decline in spontaneous activation of DMN regions with the mPFC in healthy aging and provide new evidence on the functional relevance of dynamic interactions in the human brain.

Average and transient functional interactions of the mPFC in older adults

Average CAPs reliably identified the DMN: PCC, AG, MTL and PHF all co-activated with the mPFC. Hippocampal and more extensive parahippocampal contributions were apparent in young adults. Multiple distinct patterns of brain activity contributed to this "canonical" DMN, as revealed by temporal clustering. Two out of five CAPs still strongly resembled the DMN but they differed in the involvement of limbic structures. Notably (i) the thalamus and caudate, and (ii) HC and PHF appeared in CAPs 1 and 2, respectively. This split is consistent with previous results from large groups of young healthy subjects (Liu & Duyn, 2013; Liu *et al.*, 2013). In the first study, the authors focused on the PCC, the other major DMN hub. Among 8 CAPs, the first two were most similar to the DMN and differed by the same regions as in our study. In the second study, the authors temporally clustered all scans instead of only those with supra threshold signal in a seed region. Among 30 CAPs, four resembled the DMN, which again differed by the involvement of the HC, PHF and caudate. In both studies, there were CAPs with a stronger involvement of either the HC or PHF, a subdivision we did not observe in either young or older adults. This difference could be related to the smaller number of subjects, the choice of seed, or the smaller number of CAPs we estimated. CAP 4 exclusively included superior regions of the DMN (PCC and AG) and was accompanied by a more general increase in prefrontal activity.

The other two CAPs revealed transient interactions with motor and visual cortices (CAP 5) and bilateral insulae and prefrontal areas (CAP 3). Liu & Duyn (2013) also observed co-activation of the PCC with motor and visual areas (in distinct CAPs), but, interestingly, such transient interactions were not

apparent in the exploratory whole-brain study (Liu *et al.*, 2013). Thus, our results are largely consistent with previous findings and confirm that different DMN-related CAPs occur and that, at least when using the seed-based approach, transient interactions of the DMN with other brain systems are also identified.

Age effects on functional interactions

In our study, age differences in the average CAP were mainly located in the temporal lobe (HF, PHC, MTL, temporal pole), midline parietal (Rsp) and inferior parietal cortex. This is largely consistent with previous studies, although we did not see an age effect in the posterior cingulate cortex (Andrews-Hanna *et al.*, 2007; Ferreira & Busatto, 2013).

Reduced functional interactions within the DMN have been interpreted as a disruption of the DMN in healthy aging, but our analysis of transient functional interactions suggests a different interpretation: The DMN is intact, as it can be fully recovered by temporal clustering of the fMRI data, but older adults engage one of two DMN-related CAPs less frequently than young adults. This CAP had strong contributions of regions that showed significant average effects of age (compare Figs. 6.5c and 6.6b). Two other patterns of interactions with a heavy bias for prefrontal areas were significantly overexpressed in older adults. Differences in the percentage occurrence of these three CAPs persisted after correcting for GM density and head movement during the scan. Thus, these results suggest that the disruption of average functional interactions of the DMN is – at least partly – explained by reduced and increased occurrences of different transient functional interactions.

Functional relevance of transient interactions

What is the functional significance of transient interactions of the mPFC? The mPFC is a core region of the DMN, which is thought to be involved in internally oriented cognition, such as mental simulation and self-relevant decisions (Buckner *et al.*, 2008; Andrews-Hanna *et al.*, 2010), and it is also part of a network of regions engaged in episodic memory retrieval, which includes the HC, Rsp, PCC, and AG (Vincent *et al.*, 2006; Kim, 2010; Rugg & Vilberg, 2013). The main role of the mPFC is thought to be in self-referential processes and context learning (Cabeza & St Jacques, 2007; Euston *et al.*, 2012; Rugg & Vilberg, 2013). How sensory, motor and emotional systems contribute information to these processes is not yet understood, but Buckner *et al.* (2008) suggested that transient interactions of DMN regions could serve just that purpose. Notably, in accordance with this hypothesis, the mPFC transiently interacted with both sensory and motor regions (CAP 5 in older and CAPs 3 and 5 in young adults, where we refer to Fig. E.1 for the number of the CAPs of young adults) and regions involved in salience processing (CAP 3 in older and CAPs 4 and 6 in young adults). Puzzlingly, and as already mentioned above, the PCC also transiently interacts with these systems, but whole-brain exploratory analyses did not reveal them (Liu & Duyn, 2013; Liu *et al.*, 2013).

While these transient interactions with other brain systems persist across age, their topology is modulated, as evidenced for example by the reduced co-activation of posterior DMN regions (compare CAP 3 of older and CAP 4 of young adults), and their temporal occurrence is disturbed with some transient interactions losing in importance and others gaining (as discussed above). The percentage occurrence of the frontal/insular CAP not only showed a significant effect of age, but it also significantly negatively correlated with performance in an episodic memory task in older adults. These findings

suggest a link between age-effects on dynamic network interactions and memory decline in healthy aging.

What could the functional significance of the increased occurrence of the insular/frontal CAP be? One explanation could be that the over-expression of these interactions is a trace of maladaptive over-recruitment of prefrontal areas during tasks. Over-recruitment of brain areas is commonly interpreted as a compensatory mechanism, but when not accompanied by better performance on a task, it may indicate inefficient use of neural resources or "dedifferentiation" of functional specialization (Grady, 2012). For example, increased prefrontal activity during memory encoding and retrieval correlated with worse memory performance in two recent studies (de Chastelaine *et al.*, 2011; Persson *et al.*, 2011).

Alternatively, increased spontaneous activity of this CAP could interfere with (hippocampal-dependent) learning of associations between novel input and context in the mPFC. The neocortex is generally thought to play an important role in memory only after longer delays, but Euston *et al.* (2012) for example recently suggested an early role for the mPFC in representing context and guiding behavior. Given the available data, these explanations remain speculative.

Methodological considerations

Tagliazucchi *et al.* (2012) and Liu & Duyn (2013) only selected scans with above-threshold BOLD signal in their region of interest and Liu *et al.* (2013) suggested that above and below threshold signal fluctuations contained different information. In the present study we combined the information from above- and below-threshold signal fluctuations because the average CAPs of above- and below-threshold signal were very similar and because this allowed us to approximately double the number of scans available for the clustering step. Because we focused our analysis on the CAPs of a region of interest, we could flip the sign of the scans where its BOLD signal was below threshold. This would not be possible in the whole-brain exploration of CAPs (Liu *et al.*, 2013) as all scans are retained for the clustering step. The advantage of the whole-brain analysis is that no seed region must be defined *a priori*. Here, based on extensive previous reports of a preferential anterior-posterior age-effect on the DMN and the importance of the mPFC in memory, we chose to focus our analysis on its CAPs. However, we already noted similarities between our results and the whole-brain analysis of Liu *et al.* (2013) above.

While Liu & Duyn (2013); Liu *et al.* (2013) chose the number of CAPs based on a visual comparison of results obtained with different numbers of clusters, we tackled this issue by estimating the reproducibility of the CAPs in repeated 2-fold cross-validations and by comparing it to surrogate data. This data-driven selection identified 5 reproducible CAPs and provided a more principled way to choose k. We also note that 5 appears to be a reasonable number given the available data. Of each subject's 234 scans, on average 55 were retained for temporal clustering, of which on average 13 were assigned to the most frequently occurring CAP and 8 to the least frequent one. If the number of CAPs were much larger, then even fewer scans or even zero scans for some subjects would be used for the estimation of the CAPs. Thus, the proposed cross-validation procedure appears to provide a reasonable estimate of the number of CAPs.

A further concern in the analysis of functional interactions, and even more so in the analysis of transient interactions, is head movement. While the young and old adults included in this study did not differ in root-mean-square displacement, they did differ in frame-wise displacement. To address this confound, we excluded all subjects with excessive head motion (mean displacement > 1 mm,

maximum translation or rotation > 3 mm or 3°, respectively, or mean FD > 0.5 mm); and corrected for mean FD in all analyses. Similarly, we controlled for differences in GM density in the mPFC. As our results were robust to correction for mean FD and GM, we argue that it is unlikely that our results can be explained by differences in head movement or GM volume.

Lastly, we note that we cannot exclude that some older adults suffered from preclinical Alzheimer's disease as no PiB-PET scans were available.

7 Conclusion and perspectives

FUNCTIONAL BRAIN NETWORKS estimated from fMRI data are increasingly viewed as dynamic and flexible networks, where regions take part in multiple subnetworks across time. Instead of a unique functional role, a brain region may thus have multiple functions depending on its instantaneous pattern of interactions. In this dissertation we explored multiple approaches aimed at revealing underlying structure of dynamic brain networks and understanding the functional relevance of these dynamic interactions. We summarize our main findings and mention potential future research directions.

7.1 Summary

FC pattern extraction via matrix factorizations. We have proposed to identify groups of co-varying connections by combining sliding-window correlation analysis with PCA and gCCA of a *connections* × *windows* matrix. First, by comparing PCA of dFC during rest to that of surrogate data with randomized fluctuations, we showed that groups of co-varying connections exist. The identified FC patterns revealed meaningful functional clusters of connections and their altered expression in patients with multiple sclerosis suggested a functional relevance. Second, we showed that gCCA, which emphasizes similarity across subjects, extracts similar information as PCA. Next, we compared matrix decompositions under various constraints in the temporal domain to understand whether dynamic networks are better characterized by the sole expression of a single FC pattern (temporal clustering) or the combined expression of several FC patterns (PCA). Our results from a group of young, healthy adults suggested that temporal clustering reliably identifies FC patterns, but that their joint, rather than exclusive expression, better represented dynamic networks during rest. We used 1-minute-long sliding windows for the estimation of dynamic networks and identified three different FC patterns that reliably occurred across independent split-half data sets. Our results indicated that networks involved in cognitive control contributed to changes in FC over time.

Graph wavelet transform and task-related network changes. We developed a new framework based on a tensor decomposition to define wavelets on dynamic networks. Graph wavelets respect the underlying topology of the network and are localized to a subset of the network as opposed to matrix decompositions. We showed that the tensor decomposition successfully uncovered network correlates of cognitive states in young, healthy subjects who alternated between resting and watching short movie excerpts. However, while brain graph wavelets revealed co-activations at multiple spatial

scales, this representation did not improve the detection of task-rest transitions. We have already discussed potential future research directions at the end of Chapter 5, including multi-scale analyses of brain networks (instead of brain activity) and alternative brain network definitions.

Estimation of rank-one FC patterns via temporal clustering. Because of limitations inherent to the sliding-window correlation analysis, such as a minimum window lengths of 30-60 seconds and noisy correlation estimates, we simplified the estimation of FC patterns by approximating them by rank-one matrices. This makes it possible to directly estimate these patterns from BOLD activity, which avoids the calculation of sliding-window correlations. Several authors recently proposed techniques to identify patterns of brain activity, driven by constraints in the temporal rather than spatial domain. Given the availability of promising methods, we here applied a minimally tweaked version of Liu *et al.*'s (2013) simple and easy to interpret method. We estimated CAPs by temporally clustering brain activity into distinct patterns and assessed age-related changes. Our results suggest an alternative interpretation of previous reports of disrupted functional interactions between regions of the DMN. Specifically, we demonstrated that healthy older adults have an intact DMN, but that it is engaged less frequently than in young adults. Furthermore, two CAPs occurred more frequently in older adults and one of them was associated with impaired memory.

7.2 Outlook

Alternatives to temporal clustering. K-means clustering of brain activity is attractively simple and performed very well for dFC. We here did not explore the temporal expression of the CAPs in depth, but a gradual, instead of an on-off, expression, would seem intuitive, too. Instead of "debinarizing" the weights obtained from k-means clustering as we did in Chapter 4, the non-exclusive, time-varying and non-negative expression of CAPs could also be explored by using generalizations of k-means clustering that "soft" instead of "hard" assign each data point to a cluster. Soft (or fuzzy) clustering and Gaussian Mixture Models (GMMs) assign each fMRI scan to a cluster with a strength (or probability) between 0 and 1 and each scan thus belongs to all clusters with some strength. GMM is a more formalized approach than soft clustering and (as the name suggests) models the data as a mixture of multivariate Gaussian components, which are each defined by their mean $\mu_k \in \mathbb{R}^N$ and covariance matrix $\Sigma_k \in \mathbb{R}^{N\times N}$. In k-means clustering μ_k is estimated by averaging all scans assigned to cluster k. In the "soft" approaches, μ_k and Σ_k (GMM only) are estimated after weighting each scan by its strength of belonging to component k. Note, however, that the "distance function" in GMM, one minus the probability $p(\mathbf{c}_t|k) = \frac{1}{\sqrt{(2\pi)^N \det(\Sigma_k)}} \exp\left(-\frac{1}{2}(\mathbf{c}_t-\mu_k)^T\Sigma_k^{-1}(\mathbf{c}_t-\mu_k)\right)$, where det stands for determinant, resembles the Euclidean distance more than one minus the correlation, coming with the caveats mentioned in Chapter 6. Soft clustering and GMMs, while having a weighted assignment, will still favor sparse solutions because their aim is to partition the data into several groups.

Having multiple CAPs contributing at each instant in time can also be understood as approximating FC patterns by low-rank matrices instead of rank-one matrices. Thus, more complex patterns of interactions could be modelled. These, more complex interaction patterns, are directly modelled in a GMM in Σ_k. (K-means and soft clustering model the covariance matrices of all clusters identically, that is $\Sigma_k = \sigma^2\mathbf{I}$, where $\mathbf{I}$ is the identity matrix. Even if, as an alternative, we interpreted $\mu_k \otimes \mu_k$ as rank-one approximations of network patterns in Chapter 6.) The challenge here, however, is the low temporal but high spatial resolution of fMRI. As a workaround, Eavani *et al.* (2013) parcellated the

brain into 160 regions and imposed a low-rank structure on the covariance matrices. Specifically, they modelled Σ_k as a weighted sum of a few (L) rank-one matrices: $\Sigma_k \approx \sum_{l=1}^{L} w_{kl}\mathbf{d}_l \otimes \mathbf{d}_l$, where w_{kl} is a positive weight, and $\mathbf{d}_l$ a vector (a CAP). The final goal was to identify K network states with distinct covariance structure and to achieve that they integrated the (constrained) estimation of Σ_k in a Hidden Markov Model (HMM). A HMM estimates "hidden states" (the Gaussian components) and additionally models temporal dependence (i.e., given the current state, some states are likelier to occur next than others). A network state sequence was obtained by hard assigning each fMRI scan to one state. The authors further imposed spatial sparsity on $\mathbf{d}_l$ and assumed that $L = K$ rank-one matrices serve as a basis for K unique network states. Each network state k was then defined by a unique combination $(w_{k1}, w_{k2}, \ldots, w_{kL})$ of the shared rank-one matrices. A HMM with a low-rank approximation of Σ_k is related to approximating each fMRI scan by several CAPs but is more restrictive and complex. Additionally, if we assume that multiple CAPs contribute with varying strengths over time, no explicit network states may exist (which would be in accordance with our resting-state results from Chapter 4).

Alternatively, temporal ICA has been proposed to identify groups of regions with similar temporal dynamics (Smith *et al.*, 2012). Temporal ICA, however, needs data acquired with short sampling rates and imposes temporal statistical independence, which lacks a biological foundation and may be too restrictive if there are temporal relationships among several groups of regions. Smith *et al.* (2012) suggested that relaxations of temporal independence may prove interesting, but, because they saw no evidence of sparsity in the time courses, they suggested that temporal sparsity (such as imposed by clustering or k-SVD) may not be a practical constraint. However, two observations contradict this idea. First, we and Liu & Duyn (2013) identified meaningful functional patterns with k-means clustering. Second, our simulations in Chapter 4 showed that the time courses did not directly indicate the underlying sparsity of the data because sparse data generations also led to non-sparse time courses.

When modelling dFC the positivity constraint of k-means clustering was potentially beneficial as it differentiates between FC strengths above and below average FC. When analyzing brain activity though, excursions above and below average can represent identical co-activation patterns. This is implicitly assumed by ICA and also motivated our choice to flip the signs of scans with mPFC deactivations in Chapter 6. It may thus be possible that matrix factorizations that do not constrain the time-dependent weights to be positive, or k-means clustering with one minus *absolute* correlation as a distance function, lead to interesting solutions when analyzing transient patterns of functional interactions in fMRI BOLD data.

Thus, new models that avoid the complexity and constraints of a GMM/HMM or temporal ICA, but that can explicitly and robustly model joint expression of multiple CAPs would be valuable future contributions.

Network-based alternatives. The methods discussed in the previous section impose a constraint in the time domain. Here, we discuss two alternatives that impose constraints in space, but allow more overlap than spatial ICA. The first one is "network component analysis" (Liao *et al.*, 2003). This method aims to exploit information on the network structure when solving $\mathbf{C} \approx \mathbf{DA}$, instead of imposing purely statistical constraints, such as orthogonality or independence. The algorithm starts from an approximate *a priori* network structure $\mathbf{D_0}$, which is further modified during the joint estimation of $\mathbf{D}$ and $\mathbf{A}$. As the study dynamic patterns of brain activity is only in its early days, we may however not yet be able to specify appropriate networks *a priori*. Structural networks and variants thereof may be interesting in this regard though, because they are thought to shape functional interactions.

The second approach is a variant of community detection. Communities in networks are commonly defined as groups of nodes that are particularly strongly connected to each other, but not with nodes of different communities. Such a separation into distinct groups of nodes excludes overlap. To allow overlap, Ahn *et al.* (2010) proposed to define communities as groups of connections, instead of groups of nodes. The method is based on a simple similarity measure between connections and hierarchical clustering. Schaefer *et al.* (2014) explored this approach and identified multi-network hubs, whose dFC revealed variable integration into these networks over time.

Denoising and deconvolution. FMRI data is noisy, which poses a significant challenge for all unsupervised learning methods, such as PCA, ICA or temporal clustering. Recently proposed methods that denoise the data and "undo" the hemodynamic blur to identify underlying transient events are interesting preprocessing strategies (Petridou *et al.*, 2013; Karahanoğlu *et al.*, 2013; Wu *et al.*, 2013) with initial promising results (Karahanoğlu & Van De Ville, 2014).

Subject-specific CAPs. If a_{tk} was not binary, it might also be possible to estimate subject-specific CAPs, similar to dual regression commonly employed with ICA (Beckmann *et al.*, 2009). The estimation of subject-specific spatial maps in ICA made it possible to localize changes in space by comparing the maps of different subject populations. While we could in principle also do this with binary a_{tk}, every scan would contribute to only a single CAP. Thus, every subject-specific CAP would be estimated from very few scans and the results likely very noisy. Therefore, we instead compared the temporal occurrence of each CAP, which narrows changes down to whole-brain CAPs, but not further.

7.3 Functional relevance of flexible functional interactions

Much of what we know about the significance of spontaneous *activity* and *average* functional interactions stems from examining their cognitive consequences, as reviewed in introductory Chapter 2. From a general point of view transient patterns of interactions fit well with the view of a dynamic, continuously adapting brain. The influence of prestimulus brain activity on human behavior, as for example demonstrated by Hesselmann *et al.* (2008) and Sadaghiani *et al.* (2009), already pointed to the functional relevance of instantaneous functional interactions. The study of spontaneous fluctuations in functional interactions *per se* was kicked off by the observation of strong spontaneous variability of FC within the DMN over time by Chang & Glover (2010), a finding that has since been replicated and extended to other networks. Are these fluctuations real and how do they relate to human behavior?

Two lines of evidence indicate that experimentally observed network dynamics are real and not simply a result of scanner or physiological noise or, even worse, of the analysis pipeline. First, transient interactions track experimental manipulations, as demonstrated by Gonzalez-Castillo *et al.* (2013) and in Chapters 4 and 5. Importantly, as the timing of changes in a subject's cognitive states are known to the experimenter, fluctuations in interactions can be compared to it and these studies indicate that real fluctuations can be picked up. Second, several authors have compared dynamic interactions to simultaneous physiological and electrophysiological recordings, such as the heart rate or neural activity (Chang *et al.*, 2013b,a; Allen *et al.*, 2013; Preti *et al.*, 2014). Specifically, Chang *et al.* (2013b) reported a relationship between heart rate variability and dFC of several brain regions related to arousal and vigilance. Given the enhanced concerns about noise and motion contributions for dynamic analyses, these studies provide important validations. Current results are preliminary but suggest that, at least for some dynamic interactions, physiological and neural underpinnings exist.

To better understand the functional relevance of network dynamics, several studies have aimed to assess their relationship with individual differences in human behavior. Accumulating evidence suggests that transient functional interactions are sensitive measures and can provide novel insights into psychiatric and neurological disorders and aging, as studied in Chapters 3 and 6 and reviewed in Hutchison *et al.* (2013a). Notably, altered dynamic properties may explain some of the differences observed between healthy subjects and patients in traditional, average analyses. Liu & Duyn (2013) reported that transient interactions differ between males and females and we showed in Chapter 6 that they predict individual differences in memory. Unfortunately, no dynamic behavioral correlates are currently available to which fluctuations during rest could be compared.

The purpose of spontaneous activity itself is still debated and so is that of spontaneous network reorganizations. The occurrence of dynamic interactions in anesthesized rats and monkeys (Majeed *et al.*, 2011; Hutchison *et al.*, 2013b) suggests that network dynamics are not purely a result of conscious, cognitive processing and it would be interesting to know how networks dynamics are altered in sleep or unconsciousness in humans. Sporns (2013) proposed that network dynamics are an intrinsic property of the brain, which serves to explore possible functional interactions.

7.4 Le mot de la fin

Dynamic brain networks are a relatively new topic of research. Early evidence from us and other groups suggests that dynamic analyses can provide insights into the brain's organization and its reorganization across the lifespan and in disease, which are invisible in traditional analyses. The view of the brain as a continuously changing network, rather than as a static one, is also intuitive. Much work remains to be done, however, in particular about the cognitive consequences of specific dynamic interactions and how average and dynamic analyses are best combined to complement each other. Methodological concerns persist (Handwerker *et al.*, 2012; Smith *et al.*, 2012; Hutchison *et al.*, 2013a, and see also our discussion in Appendix C). While spatial ICA and seed correlation analysis have become standard tools to study average brain networks during rest, there is currently a bewildering wealth of approaches to study network changes during rest (e.g. Kiviniemi *et al.*, 2011; Kang *et al.*, 2011; Majeed *et al.*, 2011; Hutchison *et al.*, 2013b; Liu & Duyn, 2013; Li *et al.*, 2014; Leonardi *et al.*, 2013; Eavani *et al.*, 2013; Allen *et al.*, 2014). Based on the findings in this thesis, we conclude that sliding-window correlations are a simple tool – that kicked off the field – and that can highlight whole-brain changes in FC. If data can be acquired with a short TR or if slow changes in network topology are of main interest, they can provide rich information that can be exploited with the help of multivariate statistics. If faster network reconfigurations are of interest, low-rank approximations that work directly in "activity" instead of in "connectivity" space may be more interesting, even if they constrain the network structures to be simpler. Finally, the development of standard analysis pipelines will be key to identify consistent results across different datasets and research groups.

A Supplementary material for Chapter 3

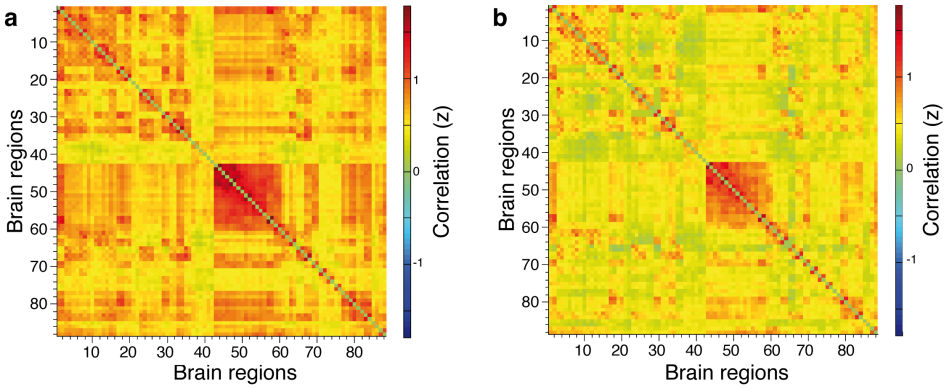


Figure A.1 – Average dynamic FC strength across time and separated for (a) HC subjects, (b) RRMS patients.

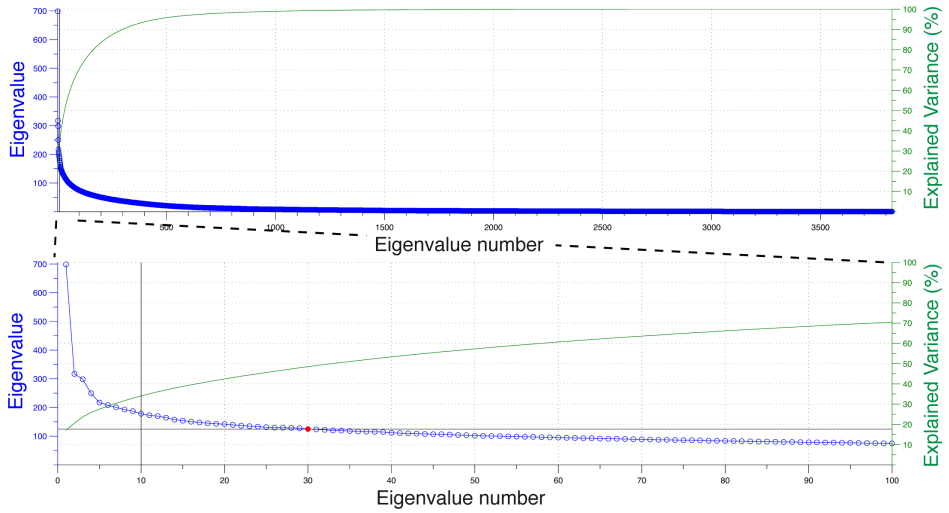


Figure A.2 – Plot of the eigenvalues $\lambda_1, \lambda_2, ..., \lambda_{3828}$ (blue), cumulative explained variance (green), 95th percentile of maximum eigenvalue in phase-randomized data ($n = 100$, corresponding to a corrected significance level of 5%), and selected cut-off based on sample size (black).

Number (L - R)	Region name	Lobe
1-2	Precentral gyrus (PreC)	Central
3-4	Superior frontal gyrus (SFG)	Frontal
5-6	Superior frontal gyrus (SFGorb), orbital	Frontal
7-8	Middle frontal gyrus (MFG)	Frontal
9-10	Middle frontal gyrus (MFGorb), orbital part	Frontal
11-12	Inferior frontal gyrus (IFGoper), opercular part	Frontal
13-14	Inferior frontal gyrus (IFGtri), triangular part	Frontal
15-16	Inferior frontal gyrus (IFGorb), orbital part	Frontal
17-18	Rolandic operculum (RolOp)	Central
19-20	Supplementary motor area (SMA)	Frontal
21-22	Olfactory cortex	Frontal
23-24	Superior frontal gyrus (SFG), medial	Frontal
25-26	Superior frontal gyrus (SFG), dorsolateral	Frontal
27-28	Gyrus rectus (Rect)	Frontal
29-30	Insula	
31-32	Anterior cingulate gyrus (ACC)	Limbic
33-34	Middle cingulate gyrus (MCC)	Limbic
35-36	Posterior cingulate gyrus (PCC)	Limbic
37-38	Hippocampus (HG)	Limbic
39-40	Parahippocampal gyrus (PAG)	Limbic
41-42	Amygdala	Limbic
43-44	Calcarine cortex	Occipital
45-46	Cuneus	Occipital
47-48	Lingual gyrus	Occipital
49-50	Superior occipital gyrus (SOG)	Occipital
51-52	Middle occipital gyrus (MOG)	Occipital
53-54	Inferior occipital gyrus (IFG)	Occipital
55-56	Fusiform gyrus	Occipital
57-58	Postcentral gyrus	Central
59-60	Superior parietal gyrus (SPG)	Parietal
61-62	Inferior parietal gyrus (IPG)	Parietal
63-64	Supramarginal gyrus (SupMarg)	Parietal
65-66	Angular gyrus (Ang)	Parietal
67-68	Precuneus (PreC)	Parietal
69-70	Paracentral lobule (ParaCen)	Frontal
71-72	Caudate	Subcortical
73-74	Putamen	Subcortical
75-76	Thalamus (Thal)	Subcortical
77-78	Heschl's gyrus	Temporal
79-80	Superior temporal gyrus (STG)	Temporal
81-82	Superior temporal pole (STP)	Temporal
83-84	Middle temporal gyrus (MTG)	Temporal
85-86	Middle temporal pole (MTP)	Temporal
87-88	Inferior temporal gyrus (ITG)	Temporal

Table A.1 – Region number (left - right), name (abbreviation) and lobe defined in the AAL atlas

Connection between ...	and ...	Cohen's d (corrected p value)
R precentral gyrus (2)	R IFG (54)	2.1 (0.01)
L mid frontal (7)	R MTP (86)	2.0 (0.02)
R IFG, triangular part (14)	L insula (29)	2.0 (0.03)
R medial SFG (24)	R rolandic operculum (18)	2.1 (0.01)
	L rectus gyrus (27)	2.0 (0.02)
L dorsolateral SFG (25)	R cuneus (46)	2.2 (0.01)
R dorsolateral SFG (26)	R MFG, orbital part (10)	1.9 (0.04)
	R SFG, medial (24)	2.4 (0.003)
	R thalamus (76)	2.6 (0.002)
L MCC (33)	R cuneus (46)	2.0 (0.04)
	L SOG (49)	2.0 (0.04)
R MCC (34)	L IFG, opercular part (11)	1.9 (0.04)
R PCC (36)	L IFG, orbital part (15)	2.0 (0.04)
L amygdala (region 41)	inferior occipital (53, 54)	2.1 (0.02), 2.8 (< 0.001)
	fusiform gyrus (55, 56)	2.4 (0.004), 2.0 (0.02)
	postcentral gyri (57, 58)	2.4 (0.005), 2.1 (0.02)
	L supramarginal gyrus (63)	2.2 (0.01)
	L angular gyrus (65)	2.0 (0.03)
	R STP (82)	2.0 (0.03)

Table A.2 – Connections with different average FC strength between HC subjects and RRMS patients

Subject	Max head motion [°/mm]	Mean FD [mm]	Mean D [mm]
1	0.70°	0.13	0.03
2	1.30 mm	0.15	0.04
3	0.89°	0.17	0.08
4	1.30 mm	0.16	0.06
5	0.77 mm	0.19	0.08
6	0.77 mm	0.11	0.04
7	0.73 mm	0.22	0.07
8	0.32 mm	0.13	0.04
9	0.57°	0.19	0.07
10	0.68 mm	0.13	0.03
11	0.28 mm	0.13	0.04
12	1.1 mm	0.16	0.05
13	1.4°	0.15	0.05
14	0.58°	0.14	0.04
15	0.66°	0.16	0.05
16	0.46 mm	0.11	0.04
17	0.56 mm	0.10	0.03
18	0.68 mm	0.12	0.03
19	0.28 mm	0.20	0.04
20	0.50°	0.16	0.06
21	0.74 mm	0.12	0.05
22	0.36 mm	0.24	0.07
23	0.58°	0.13	0.03
24	0.24 mm	0.14	0.02
25	1.50 mm	0.22	0.09
26	0.87 mm	0.23	0.09
27	0.77 mm	0.10	0.04
28	0.42°	0.12	0.06

Table A.3 – Summary of subject motion. The second column lists the largest head motion parameter among the 6 parameters. Note that the first column lists absolute displacement while columns 2 and 3 list relative displacement between adjacent volumes.

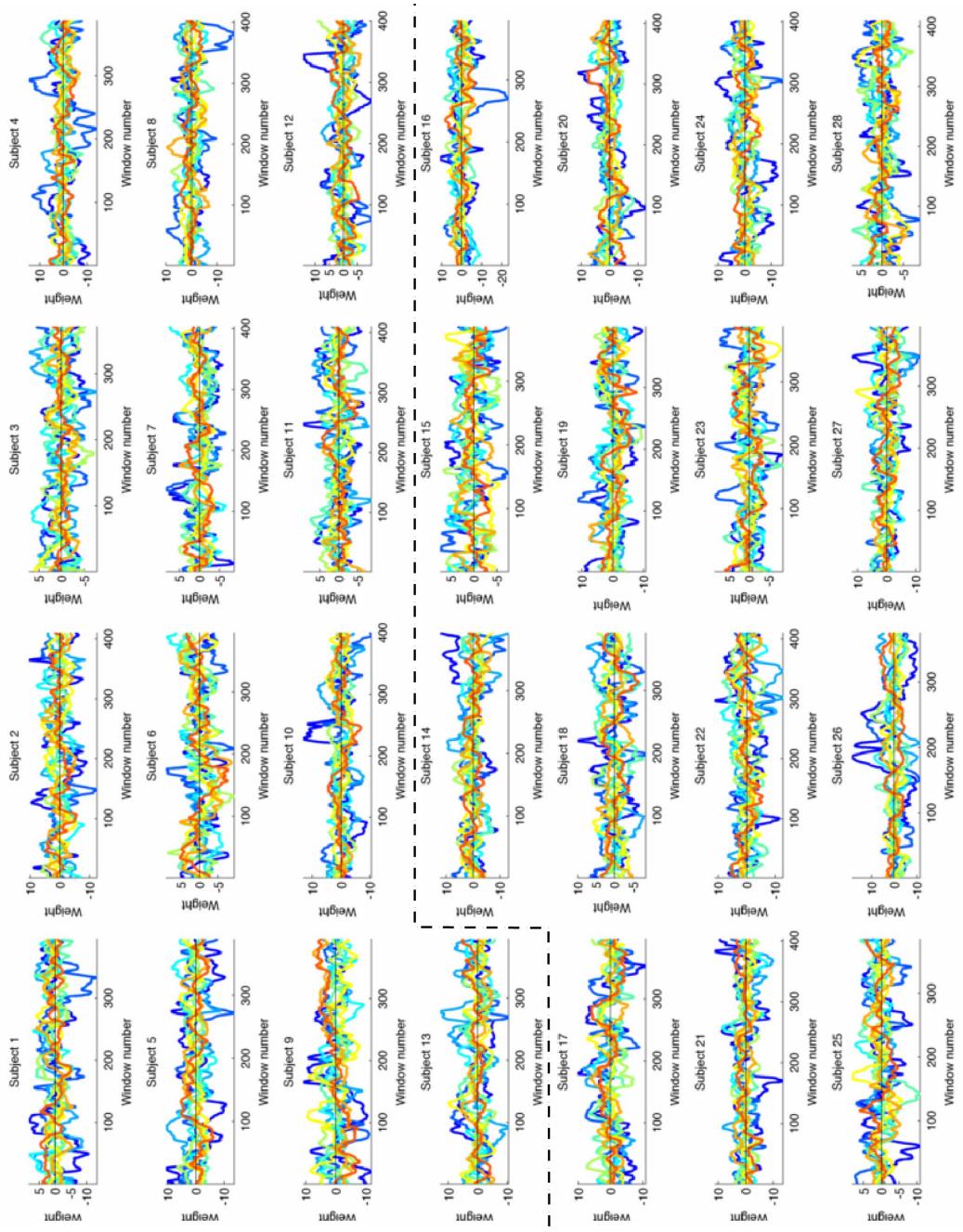


Figure A.3 – Time-dependent weights of eigenconnectivities 2 – 10 for all subjects. Weights of eigenconnectivity 1 are not shown because it has a larger range than the others. Subject 6 is the one shown in the main manuscript, subjects 1-13 are HC subjects, 14-28 RRMS patients (separated by dashed line).

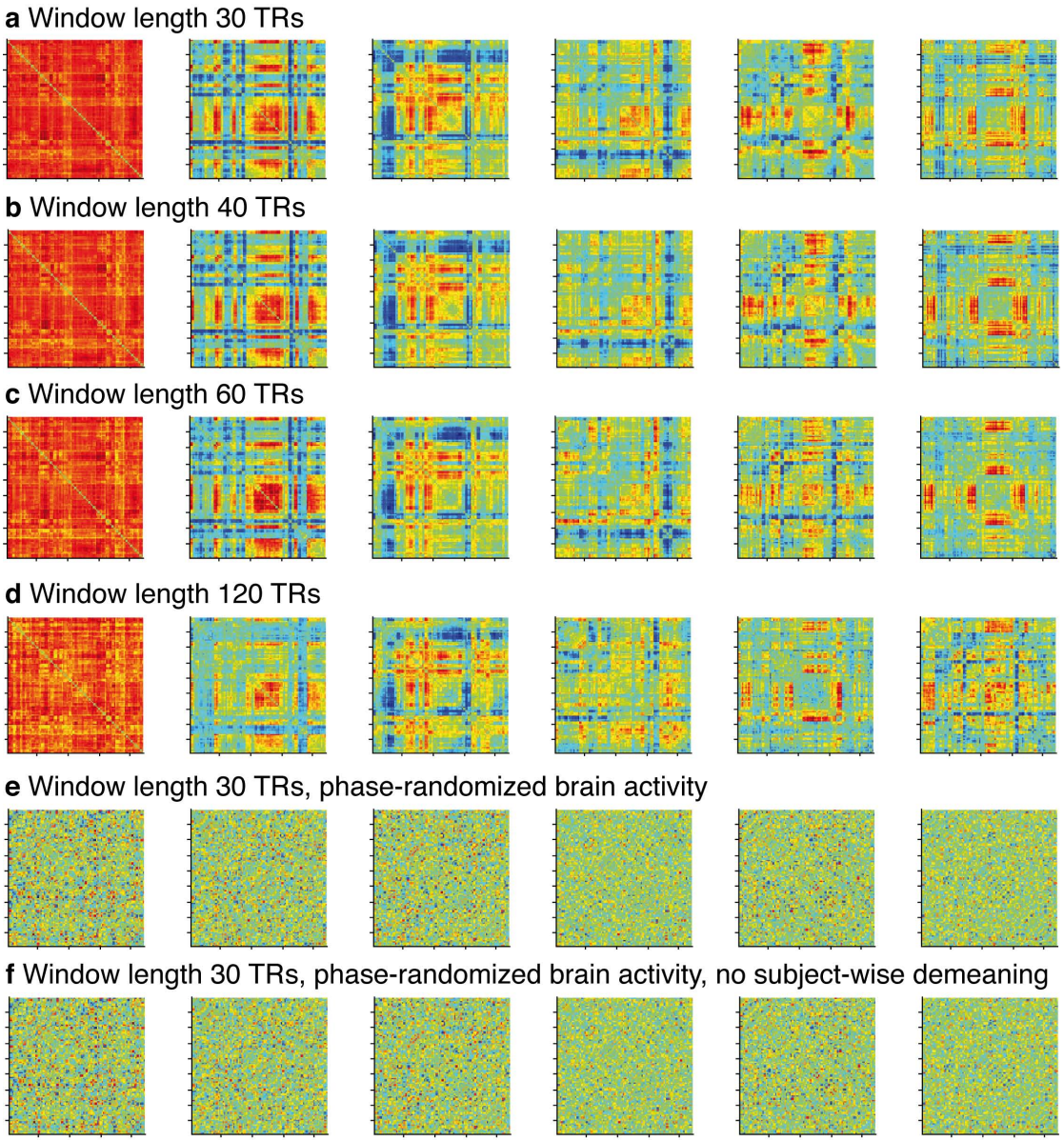


Figure A.4 – (a-d) Eigenconnectivities estimated using different window lengths. Note that the sign is essentially arbitrary, i.e., similar connectivity patterns that change from red to blue color for different window lengths encode similar information. (e-f) Eigenconnectivities estimated from phase randomized activity time courses with (e) and without (f) row-wise demeaning of each subject's dynamic FC matrix. Data matrix concatenated across time- and subject as for the original data. Colorbars are symmetric around zero

B Supplementary material for Chapter 4

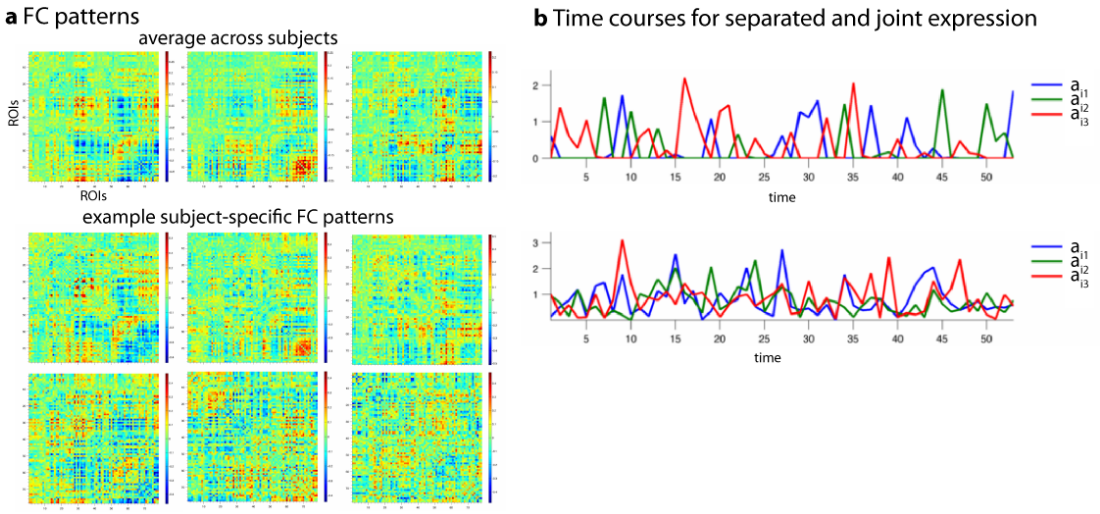


Figure B.1 – (a) FC patterns used to generate the simulated data, averaged across subjects and for a "good" and "bad" subject. (b) Time courses with sparsity parameters S = 1 or S = 3 for an example subject. Final data for each subject was generated by multiplying the vectorized patterns with their associated time courses and adding Gaussian noise.

C Spurious and real fluctuations in sliding-window correlations

C.1 Journal article

Submitted to: NeuroImage

On spurious and real fluctuations of dynamic functional connectivity during rest

Nora Leonardi,[1,2] and Dimitri Van De Ville[1,2]

[1] Medical Image Processing Lab, Institute of Bioengineering, École Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland; [2] Medical Image Processing Lab, Department of Radiology and Medical Informatics, University of Geneva, Geneva, Switzerland

Functional brain networks reconfigure spontaneously during rest. Such network dynamics can be studied by dynamic functional connectivity (dynFC); i.e., sliding-window correlations between regional brain activity. Key parameters—such as window length and cut-off frequencies for filtering—are not yet systematically studied. In this letter we provide the fundamental theory from signal processing to address these parameter choices when estimating and interpreting dynFC. We guide the reader through several illustrative cases, both simple analytical models and experimental fMRI BOLD data. First, we show how spurious fluctuations in dynFC can arise due to the estimation method when the window length is shorter than the largest wavelength present in both signals, even for deterministic signals with a fixed relationship. Second, we study how real fluctuations of dynFC can be explained using a frequency-based view, which is particularly instructive for signals with multiple frequency components such as fMRI BOLD, demonstrates that fluctuations in sliding-window correlation emerge by interaction between frequency components similar to the phenomenon of beat frequencies. We conclude with practical guidelines for the choice and impact of the window length.

Introduction

Functional magnetic resonance imaging (fMRI) has become a key tool to probe the large-scale organization of the brain. Functional connectivity (FC), which is estimated by correlation of BOLD

activity, identifies coherent brain activity in distributed and reproducible networks. FC has revealed reorganization of brain networks during cognitive tasks (Lewis *et al.*, 2009; Richiardi *et al.*, 2011; Shirer *et al.*, 2012; Ekman *et al.*, 2012), but also at rest (Chang & Glover, 2010; Majeed *et al.*, 2011; Kang *et al.*, 2011; Allen *et al.*, 2014; Smith *et al.*, 2012; Hutchison *et al.*, 2013b; Leonardi & Van De Ville, 2013b). To study changes in FC over time sliding-window correlation analysis, where the correlation is estimated for brain activity during multiple, possibly overlapping temporal segments (typically 30-60 seconds), has been widely deployed (Chang & Glover, 2010; Allen *et al.*, 2014; Hutchison *et al.*, 2013a). A caveat of analyzing dynamic FC (dFC) by sliding-window correlation is that the small number of time points renders the estimates unreliable and might lead to spurious variability of dFC (Smith *et al.*, 2012; Hutchison *et al.*, 2013a). However, there is no systematic account that perspeciously indicates the trade-off that is made by choosing the window length, and its implications for filtering of BOLD activity time series and dFC itself.

We first break sliding-window correlation into several components to facilitate its study. Then, we present a simple yet instructive analytical model to study the emergence of spurious variability of dFC in stationary signals. In particular, we investigate the influence of various parameters such as frequency, phase lag, and window length. Next, we introduce a small change to our analytical model to study how real variability of dynFC due to non-stationarity might arise. To provide the best possible insights for signals with many frequency components, we present a frequency-based view on dynFC. This provides an elegant explanation of how fluctuations of dynFC emerge through the interaction between different frequency components. Finally, we illustrate dynFC between two main regions of the default-mode network with experimental fMRI data.

Breaking down sliding-window correlations

We start by reformulating sliding-window correlation into simpler terms. In particular, we first look at sliding-window *covariance*, which for two time series x and y with sampling period TR is defined as follows at scan n:

$$c_{xy}[n] \quad = \quad \mathrm{cov}(x[n-\Delta, n+\Delta], y[n-\Delta, n+\Delta]) = \frac{\mathrm{TR}}{w} \sum_{i=n-\Delta}^{n+\Delta} (x_i - \bar{x}_n)(y_i - \bar{y}_n), \tag{C.1}$$

where $w = (2\Delta + 1)\mathrm{TR}$ is the odd window length in seconds, i sums only over the scans inside the window, and

$$\bar{x}_n = \frac{\mathrm{TR}}{w} \sum_{i=n-\Delta}^{n+\Delta} x_i$$

is the local average inside the window at position n. This calculation is then repeated for all values of n ("sliding" the window across time). After some elementary manipulations, we arrive at the following equality:

$$\begin{aligned} c_{xy}[n] \quad &= \quad \frac{\mathrm{TR}}{w} \sum_{i=n-\Delta}^{n+\Delta} (x_i - \bar{x}_n)(y_i - \bar{y}_n) = \frac{\mathrm{TR}}{w} \sum_{i=n-\Delta}^{n+\Delta} x_i (y_i - \bar{y}_n) - \mathrm{TR}\, \frac{\bar{x}}{w} \sum_{i=n-\Delta}^{n+\Delta} (y_i - \bar{y}_n) \\ &\overset{(a)}{=} \quad \frac{\mathrm{TR}}{w} \sum_{i=n-\Delta}^{n+\Delta} x_i (y_i - \bar{y}_n) = \frac{\mathrm{TR}}{w} \sum_{i=n-\Delta}^{n+\Delta} x_i y_i - \bar{y}_n \frac{\mathrm{TR}}{w} \sum_{i=n-\Delta}^{n+\Delta} x_i \\ &= \quad \underbrace{\frac{\mathrm{TR}}{w} \sum_{i=n-\Delta}^{n+\Delta} x_i y_i}_{\mathrm{I}} - \underbrace{\bar{y}_n \bar{x}_n}_{\mathrm{II}}, \end{aligned} \tag{C.2}$$

where $^{(a)}$ simplifies as the second term equals zero. Thus, $c_{xy}[n]$ can be separated into two terms, which are the local average of the cross-product xy (I) minus the product of the local averages of x and y (II).

The sliding-window correlation is then obtained by normalizing at each window by the local variances:

$$\rho_{xy}[n] = \frac{c_{xy}[n]}{\sqrt{c_{xx}[n]c_{yy}[n]}}. \tag{C.3}$$

Having identified the components that constitute sliding-window correlation, we can now analyze and understand dFC more easily.

Spurious fluctuations in dFC

Effect of the window length

We want to understand how spurious fluctuations of dynFC might arise even for deterministic signals with a fixed relationship; i.e., we consider two pure sinusoidal signals that are phase-locked. Specifically, we take

$$x_i = \sqrt{2}\cos(2\pi f i\ \mathrm{TR}), \quad y_i = \sqrt{2}\cos(2\pi f i\ \mathrm{TR} + \theta), \tag{C.4}$$

where the factor $\sqrt{2}$ normalizes both signals for variance equal to one per time unit. This normalization makes the sliding-window covariance comparable to sliding-window correlation as a first approximation; i.e., we have the asymptotic equivalence $\lim_{w\to\infty}\rho_{xy}[n] = c_{xy}[n]$.

To investigate the influence of the key parameters frequency f, phase lag θ, and window length w, we derive the analytical form of $c_{xy}[n]$ for the signals of Eq. (C.4). First, we approximate $\bar{y}_n$ by integration as follows:

$$\begin{aligned}
\bar{y}_n &= \frac{\mathrm{TR}}{w}\sum_{i=n-\Delta}^{n+\Delta}\sqrt{2}\cos(2\pi f i\ \mathrm{TR}+\theta) \\
&\approx \frac{\sqrt{2}}{w}\int_{(n-\Delta)\mathrm{TR}}^{(n+\Delta)\mathrm{TR}}\cos(2\pi f t+\theta)\,dt \\
&= \frac{\sqrt{2}}{w}\left[\frac{1}{2\pi f}\sin(2\pi f t+\theta)\right]_{(n-\Delta)\mathrm{TR}}^{(n+\Delta)\mathrm{TR}} \\
&= \frac{\sqrt{2}}{w2\pi f}\big(\sin(2\pi f(n+\Delta)\mathrm{TR}+\theta) - \\
&\quad \sin(2\pi f(n-\Delta)\mathrm{TR}+\theta)\big) \\
&= \frac{\sqrt{2}}{w\pi f}\cos(2\pi f n\mathrm{TR}+\theta)\sin(2\pi f\Delta\mathrm{TR}).
\end{aligned}$$

On similar grounds, we also find $\bar{x}_n = \frac{\sqrt{2}}{w\pi f}\cos(2\pi f n\mathrm{TR})\sin(2\pi f\Delta\mathrm{TR})$. Therefore, the second term $\bar{x}_n\bar{y}_n$ of Eq. (C.2) reverts to

$$\bar{x}_n\bar{y}_n = \frac{2}{w^2\pi^2 f^2}\cos(2\pi f n\mathrm{TR})\cos(2\pi f n\mathrm{TR}+\theta)\sin^2(2\pi f\Delta\mathrm{TR}).$$

To estimate the first term of Eq. (C.2), we use the product-to-sum trigonometric identity

$$2\cos(2\pi f i\mathrm{TR})\cos(2\pi f i\mathrm{TR}+\theta) = \cos(4\pi f i\mathrm{TR}+\theta)+\cos(\theta),$$

which, after integration, leads to

$$\cos(\theta)+\frac{1}{w\pi f}\cos(2\pi f n\mathrm{TR}+\theta)\sin(2\pi f\Delta\mathrm{TR}).$$

By combining both terms, we retrieve the expression

$$\begin{aligned} c_{xy}[n] &= \cos(\theta)+\frac{1}{w\pi f}\cos(2\pi f n\mathrm{TR}+\theta)\sin(2\pi f\Delta\mathrm{TR}) - \\ &\quad \frac{2}{w^2\pi^2 f^2}\cos(2\pi f n\mathrm{TR})\cos(2\pi f n\mathrm{TR}+\theta)\sin^2(2\pi f\Delta\mathrm{TR}). \end{aligned} \tag{C.5}$$

As a sanity check, we see that in the limit of stationary covariance (i.e., infinite window length), we have

$$\lim_{w\to+\infty} c_{xy}[n] = \cos(\theta).$$

We now use this expression to efficiently trace $c_{xy}[n]$ as function of frequency f, phase lag θ, window length w, and window position n. In Fig. C.1a, $c_{xy}[n]$ is plotted for $f = 0.025$Hz and zero phase lag, as a function of window length w. The dashed lines are for different window positions n, and the thick line corresponds to the mean $\bar{c}_{xy} = E[c_{xy}[n]]$. We observe considerable fluctuations of $c_{xy}[n]$ for short window lengths, and crossings with the true value (i.e., 1) exactly for multiples of the window length because the term $\sin(2\pi f\Delta\mathrm{TR})$ in Eq. (C.5) vanishes for $2\Delta\mathrm{TR} = 1/f$. Importantly, only when the window length is larger than the first crossing, which corresponds to the wavelength $1/f = 40$s, fluctuations of $c_{xy}[n]$ diminish and converge to the true value of $\cos(\theta)$.

The same observations can be made from Figs. C.1b and C.1c, where we plot $\bar{c}_{xy}$ for various frequencies, and the difference between maximal and minimal $c_{xy}[n]$ in Fig.C.1d. Spurious fluctuations of $c_{xy}[n]$ occur when the window length is too short with respect to the underlying frequency component. We propose the following rule of thumb for minimal window length when observing underlying frequencies of $f_{\min}$ or higher:

$$w \geq \frac{1}{f_{\min}}.$$

Therefore, high-pass filtering that removes frequency components below $1/w$ can be recommended; see also Smith *et al.* (2012) and Hutchison *et al.* (2013a) for similar recommendations. The cut-off frequency $f_{\min}$ is indicated in Fig. C.1. It should be noted that these plots only depend on the window length in seconds, not in TRs.

Sliding-window correlation ρ_{xy} (and its fluctuations) can be obtained by normalizing $c_{xy}[n]$ according to Eq. (C.3). In the ideal case with zero phase lag, sliding-window correlation clamps to 1, however, even a small phase lag is sufficient to introduce the same spurious fluctuations as we observed for sliding-window covariance. In Fig. C.2, we plot sliding-window correlation and its extrema for phase lags of $\theta = \pi/16$ and $\theta = \pi/4$ respectively. The variability of sliding-window correlation is decreased compared to sliding-window covariance and the true correlation of $\cos(\theta)$ is recovered for window lengths above $w_{\min}$, in accordance with the previous rule of thumb.

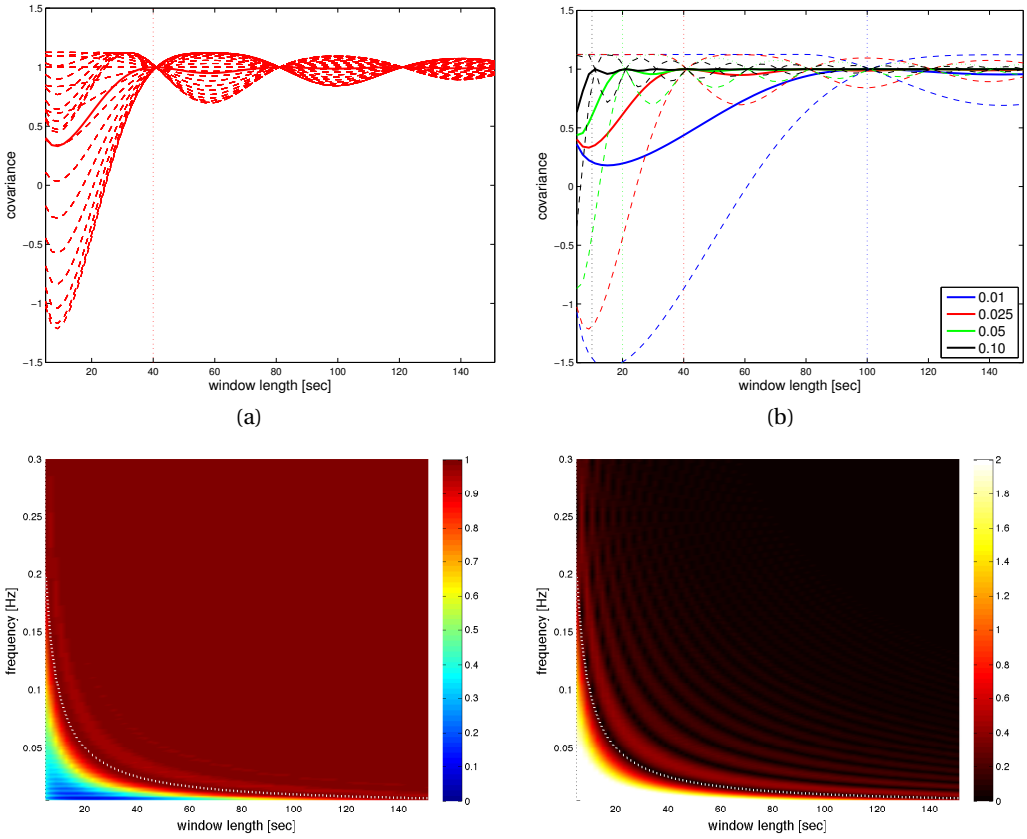


Figure C.1 – Sliding-window covariance $c_{xy}[n]$ for two pure sinusoids without phase lag; i.e., the true covariance is always 1. (a) For frequency $f = 0.025$Hz as a function of window length. Thick line indicates the mean $\bar{c}_{xy}$ over all shifts n of the window position. Dashed lines indicate $c_{xy}[n]$ for different n. (b) For different frequencies $f = 0.01, 0.025, 0.05, 0.10$Hz as a function of window length. Thick line indicates $\bar{c}_{xy}$, dashed lines $\min_n c_{xy}[n]$ and $\max_n c_{xy}[n]$, and vertical lines the minimal window lengths $w = 1/f_{\min}$. (c) Landscape of $\bar{c}_{xy}$ as a function of frequency and window length. The dashed line indicates $w = 1/f_{\min}$. (d) Landscape of maximum – minimum of $c_{xy}[n]$ as a function of frequency and window length.

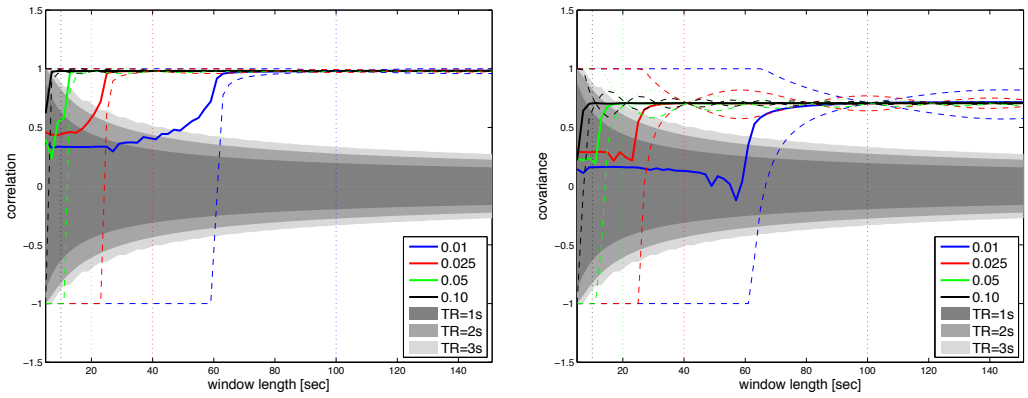
Effect of sampling and noise

Up to now, we did not take into account the effect of sampling or noise contributions. For the estimation of sliding-window correlation, we can use the 5% confidence interval for the standard test of significant non-zero correlation:

$$\rho^* = \frac{t}{\sqrt{w/\mathrm{TR} - 2 + t^2}},$$

where t follows Student's t-distribution with $w/\mathrm{TR} - 2$ degrees of freedom. The confidence intervals for several TRs are shown in Fig. C.2b and we list ρ^* for common choices of w and TR in the Appendix. Clearly, the number of scans included inside the window is an important factor for the estimator and thus depends on the TR for fixed window length (in seconds). Alternatively, phase randomization of the original timecourses could serve as a non-parametric approach to derive surrogate data with the same autocorrelation properties (Handwerker *et al.*, 2012).

In sum, we have shown that spurious fluctuations of dynFC can arise by analyzing components



(a) (b)

Figure C.2 – (a) Sliding-window correlation $\rho_{xy}[n]$ for two pure sinusoids with phase lag $\pi/16$. (b) Sliding-window correlation $\rho_{xy}[n]$ for phase lag $\pi/4$. The gray shaded area indicates the 5% confidence interval of significant correlation for different TRs=1,2,3s.

with wavelengths larger than the window length, and that the limited number of data points inflates the influence of noise on the correlation estimates. The latter effect has also been referred to as "poor sampling of correlation" (Smith *et al.*, 2012).

Real fluctuations in dFC due to non-stationarity

Effect of modulatory component

We now first make a slight modification to the analytical model to introduce non-stationary relationship between both signals and thus true variability in dynFC. Specifically, we modify y by multiplying it with a low-frequency component ($f_0 << f$),

$$y_i = \sqrt{2}\cos(2\pi f\, i\mathrm{TR})\cos(2\pi f_0\, i\mathrm{TR}).$$

Then, dFC between x and y will vary between $+1$ and -1 depending on the phase with respect to the low-frequency component. We note that the signal y is equivalent to

$$y_i = \frac{\sqrt{2}}{2}\left(\cos(2\pi(f+f_0)\, i\mathrm{TR}) + \cos(2\pi(f-f_0)\, i\mathrm{TR})\right),$$

which shows that low-frequency modulation is equivalent to introducing more frequency components. Repeating the analytical derivation of the previous section is possible, but the formulas rapidly become involved due to the interaction of many sinusoids. Instead, we revisit the sliding-window covariance of Eq. (C.2) and provide a Fourier interpretation that shows more easily how interactions between frequency components can contribute to dFC. Such a frequency-based view of dFC is particularly helpful for real fMRI data with broad spectra.

The repeated calculation of local averages by shifting the window can be written as the convolution with a rectangular window h

$$c_{xy} = (xy) * h - (x * h)(y * h), \tag{C.6}$$

where $*$ denotes convolution and h is defined as

$$h[n] = \frac{\mathrm{TR}}{w}\mathrm{rect}\left(\frac{n\mathrm{TR}}{w}\right) = \begin{cases} \frac{\mathrm{TR}}{w}, & \text{for } |n| \le \Delta, \\ 0, & \text{otherwise.} \end{cases}$$

For a tapered window, the local sums become weighted sums.

Using the convolution theorem[1], the discrete Fourier transform (DFT) of Eq. (C.6) can be obtained as

$$C_{xy} = (X * Y)H - (XH) * (YH), \tag{C.7}$$

where capital letters denote the discrete Fourier transform (DFT) of the signals, in particular:

$$X[k] = \mathscr{F}\{x[n]\} = \sum_{n=0}^{N-1} x[n] \exp\left(-j2\pi\frac{kn}{N}\right), \quad k = 0, 1, \ldots, N-1,$$

where N is the full length of the signal. For a signal sampled with period TR, the corresponding frequencies in Hz are given by $f_k = k/(N\mathrm{TR})$.

The DFT of the rectangular window h is well-known to be the Dirichlet kernel

$$H[k] = \frac{\mathrm{TR}}{w}\frac{\sin(\pi k w/(N\mathrm{TR}))}{\sin(\pi k/(N\mathrm{TR}))},$$

which can be seen as the discrete version of the common sinc-function. Fig. C.3a shows the window function and Fig. C.3b its spectrum, where the width of the main lobe is $1/w$ [Hz]. The convolution of x, y, and xy with h can thus be seen as low-pass filtering operations.

The DFT of our signals x and y can be obtained as[2]

$$\begin{aligned} X[k] &= \frac{\sqrt{2}}{2}\left(\delta_{fN\mathrm{TR}}[k] + \delta_{-fN\mathrm{TR}}[k]\right), \\ Y[k] &= \frac{\sqrt{2}}{4}\left(\delta_{(f+f_0)N\mathrm{TR}}[k] + \delta_{-(f+f_0)N\mathrm{TR}}[k] + \delta_{(f-f_0)N\mathrm{TR}}[k] + \delta_{-(f-f_0)N\mathrm{TR}}[k]\right), \end{aligned}$$

where δ is the Kronecker-delta function (i.e., $\delta_0[k] = 1$ for $k = 0$, and 0 otherwise). Exemplary signals x and y and their frequency spectra are shown in Figs. C.3c and d. Since we deal with real-valued signals, all amplitude spectra are Hermitian symmetric and we only depict positive frequencies in the plots. The convolution $X * Y$ then redistributes the delta functions as

$$\begin{aligned} (X * Y)[k] = {} & \frac{1}{2}\left(\delta_{f_0N\mathrm{TR}}[k] + \delta_{-f_0N\mathrm{TR}}[k]\right) + \frac{1}{4}\left(\delta_{(2f-f_0)N\mathrm{TR}}[k] + \delta_{(2f+f_0)N\mathrm{TR}}[k] + \right. \\ & \left. \delta_{(-2f-f_0)N\mathrm{TR}}[k] + \delta_{(-2f+f_0)N\mathrm{TR}}[k]\right), \end{aligned}$$

which is illustrated in Fig. C.3f. Assuming that the window length has been chosen according to the rule of thumb (i.e., we have $w > 1/f_{\min}$ and thus also $f > f_{\min}$), frequency components at $\pm 2f \pm f_0$ are well suppressed by the filtering operation $(X * Y)H$, as well as those at $\pm f$ and $\pm f \pm f_0$ in XH and YH, respectively. Consequently, the second term of Eq. (C.7) vanishes and the first term simplifies, leading

[1] Convolution in the time domain corresponds to multiplication in the frequency domain and vice versa.
[2] Without any loss of generality, we have assumed here that $fN\mathrm{TR}$ corresponds to an integer number.

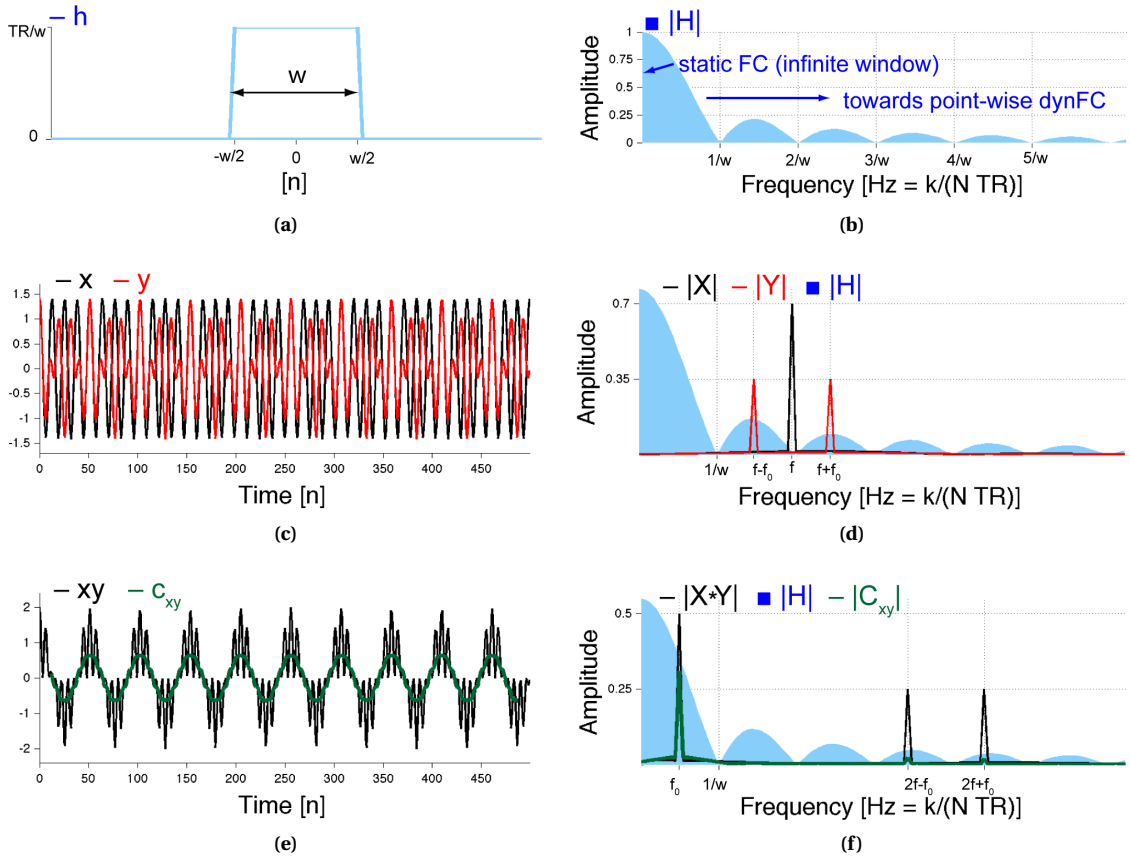


Figure C.3 – Window h in time (a) and its frequency spectrum (b). The frequency selection of the filter is governed by $1/w$. Two cosines x and y, where y is modulated by a low-frequency cosine at f_0, in time (b) and in frequency domain (c). The modulation of y corresponds to two slightly shifted peaks in Y at $f - f_0$ and $f + f_0$. The peaks at $f - f_0$, f and $f + f_0$ will be filtered out by H as they beyond the main side lobe of the filter H. The point-wise product xy and sliding-window covariance c_{xy} in time (e) and frequency (f). Three new peaks appear in the convolution $X * Y$ from the interaction of the frequency components: f_0, $2f - f_0$,$2f + f_0$. Only the peak at f_0 is retained by the sliding-window covariance.

to

$$C_{xy}[k] = \frac{\mathrm{TR}}{w}\frac{\sin(\pi f_0 w)}{\sin(\pi f_0)}\left(\frac{\delta_{f_0 N\mathrm{TR}}[k] + \delta_{-f_0 N\mathrm{TR}}[k]}{2}\right),$$

where only $k = f_0 N\mathrm{TR}$ survives. The remaining frequency component at f_0 is also illustrated in Fig. C.3d. C_{xy} can be identified as the DFT of

$$c_{xy}[n] = \frac{\mathrm{TR}}{w}\frac{\sin(\pi f_0 w)}{\sin(\pi f_0)}\cos(2\pi f_0 n\mathrm{TR}),$$

shown in Fig. C.3e. Because fluctuations in xy are low-pass filtered by the convolution with h with cut-off frequency $1/w = f_{\min}$, the slow modulation term—which in this case is a true fluctuation of dynFC—is recovered as long as $f_0 < f_{\min}$ and $f - f_0 \approx f > f_{\min}$. The influence of the window length on its low-pass filtering effect has previously been noted by Handwerker *et al.* (2012) and less variable dFC with longer windows is a well documented empirical observation (e.g., Chang & Glover, 2010;

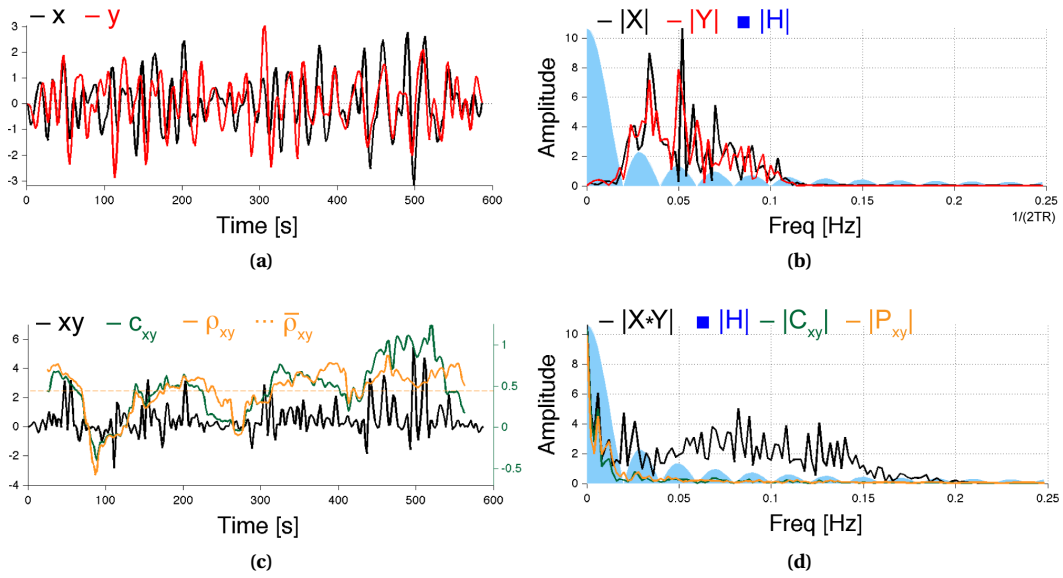


Figure C.4 – (a-b) PCC and left AG time series and their frequency spectra. (c-d) Illustration of sliding-window covariance and correlation estimation in time and frequency. xy is the point-wise multiplication of the two time series, c_{xy} the sliding-window covariance and ρ_{xy} the sliding-window correlation. Note that the Fourier spectra of c_{xy} and ρ_{xy} are similar. $\bar{\rho}_{xy}$ is the average of dFC across time, corresponding to a static FC estimate.

Hutchison *et al.*, 2013b; Leonardi & Van De Ville, 2013b). The spectral selectivity of the windowing operation can be improved by using tapering; e.g., Hamming filter (Handwerker *et al.*, 2012), Gaussian filter (Allen *et al.*, 2014), or other windows with smooth roll-off at the edges (Smith *et al.*, 2012). In such case, the window length should be replaced by the equivalent window length that corresponds to the cut-off frequency of the tapered window. It is essential to note that the frequency component f_0 emerges by "interaction" between both spectra and is not present as such in the original spectra. Mathematically, the beat frequency is recovered by multiplication in the time-domain, or, equivalently, convolution in the Fourier domain.

Example of experimental fMRI data

The frequency-based view is particularly instructive for experimental BOLD data because they have broad spectra that are not easily understood in terms of single frequency components. We illustrate dFC between two key regions of the default-mode network: the posterior cingulate cortex (PCC) and left angular gyrus (AG). Changes of FC over time between these regions have been previously demonstrated (e.g. Chang & Glover, 2010). Two regionally-averaged time series were extracted from a 10-minute long resting-state fMRI scan (data acquisition and preprocessing as in Shirer *et al.*, 2012). We choose the window length to be $w = 50$ sec, or 25 scans. Consequently, we high-pass filtered the time series with a cut-off frequency of $1/w = 0.02$ Hz, see Figs. C.4a and b. Next, the point-wise multiplication of both time series and c_{xy} (~ low-pass filtered xy) are computed, shown in Fig. C.4c. Clearly, these time series are similar during a majority of the scan, but their similarity is strongly diminished during two periods in time (around 90 and 270 sec) and another small dip is visible at 420 sec. In the Fourier domain, shown in Fig. C.4d, this variability is apparent from new low-frequency components, notably a peak at 0.006 Hz, which appears to be approximately the wavelength of the interaction cycles. Because of

the broad spectra of X and Y *multiple* frequency components contribute to these new low-frequency components.

In Figs. C.4c and d, we also show sliding-window correlation ρ_{xy} in time and frequency. While the normalization reduces some fluctuations, its main characteristics are the same as for c_{xy} and are determined by $(X * Y)H$.

Finally, it is important to note that while fluctuations of dFC might be due to true nonstationarities and interactions of the time series, the origin of these signals could be both neurological and non-neurological. Therefore, future research that validates fluctuations of dFC using concurrent measurements, such as electroencephalography (EEG), will likely provide further insights.

Conclusion

Taking advantage of an analytical model, we have derived some important properties of dynFC that explain the emergence of spurious fluctuations due to a mismatch between the choice of the window length and high-pass filtering of the original timecourses, as well as how potentially real fluctuations can arise due to modulatory components. We conclude that the window length w, specified in seconds, is the key parameter that needs to be chosen carefully as it sets following trade-offs:

1. Spurious fluctuations of dynFC due to the nature of the estimation method are limited by high-pass filtering of the original time series with cut-off frequency $1/w$.

2. Remaining fluctuations of dynFC are low-pass filtered with cut-off frequency $1/w$.

3. Variability due to the influence of noise on the estimator increases with smaller window lengths or longer TR.

These different criteria should be balanced well. For example, windows as short as 20s would require the removal of low frequency components up to 0.05 Hz, which are typically of interest in resting-state studies, and the confidence interval for significant correlations would be very high ($\rho > 0.63$ for a TR of 2 seconds). Given these considerations typical choices of window lengths (30-60 seconds) appear reasonable and these lengths have also been backed up by empirical studies that discriminated between cognitive states (Shirer *et al.*, 2012; Gonzalez-Castillo *et al.*, 2013). When interpreting dFC spectra, we suggest to focus on the frequency interval $[0 - 1/w]$ Hz because of the low-pass filtering effect of the window on dFC, which means that the modulatory effects that can be observed are relatively slow; e.g., up to 0.16 Hz for a 60 sec window length.

One promising future avenue to overcome the choice of a fixed window length, is the use of the wavelet transform, which would allow to conveniently focus on particular frequencies (scales). In particular, the wavelet transform coherence (WTC) has been suggested as one alternative to estimate instantaneous "correlation coefficients" at different frequency bands, with a window length adjusted to the frequency content of the signal (Torrence & Webster, 1999; Chang & Glover, 2010; Hutchison *et al.*, 2013a). In particular, the signal at each wavelet scale has been band-pass filtered with a high-pass cut-off according to the rule of the thumb, and, consequently, temporal variations in the scale-dependent correlations are limited to the same frequency. While this gives access to a rich amount of information, one remaining issue is how to combine scales into a concise and meaningful summarizing measure.

Finally, it is important to note that while fluctuations of dynFC might be driven by true non-stationarities and interactions of the time series, the origin of these signals could be both neurological

and non-neurological. Ongoing and future research should further validate to what extend these origins can be disentangled (Chang *et al.*, 2013a); e.g., using concurrent measurements such as electroencephalography (EEG) and non-neurophysiological signals.

Appendix

	TR		
Window length	1s	2s	3s
20s	0.44	0.63	0.75
30s	0.36	0.51	0.63
40s	0.31	0.44	0.55
50s	0.28	0.40	0.48
60s	0.25	0.36	0.44
120s	0.18	0.25	0.31

Table C.1 – ρ^* for typical window lengths and TRs (w/TR was rounded to the nearest integer).

D Connectivity-based decoding of cognitive states

We reproduce the poster presented at the 2012 Annual Meeting of the Organization for Human Brain Mapping (OHBM), where we compared **activity- and connectivity-based approaches** for the classification of individual fMRI scans into one of two cognitive states:

1. Activity-based: We use regional brain activity as our feature to predict the cognitive state (referred to as *baseline* on the poster),
2. Connectivity-based (global): We use the projection of brain activity onto the eigenvectors of the graph Laplacian matrix of each cognitive state as our feature (*Fourier*),
3. Connectivity-based (local and multi-scale): We use the analysis of brain activity with graph wavelets adapted to the FC of each state as our feature (*wavelets*).

Additional information on methods

To compare the various approaches, we used both simulated and experimental fMRI data. For the simulations, we generated 3 synthetic datasets where (i) only activity levels changed between the two conditions ($\mathbf{x} \sim \mathcal{N}(\mu_i, \mathbf{I})$, where $\mathbf{I}$ is the identity matrix and $i = 1,2$), (ii) only the covariance changed ($\mathbf{x} \sim \mathcal{N}(0, \Sigma_i)$), and (iii) both changed ($\mathbf{x} \sim \mathcal{N}(\mu_i, \Sigma_i)$). We used an approach similar to the one described in Ng *et al.* (2011) to simulate the data. The experimental fMRI data and preprocessing steps are as described in the journal article reproduced in Chapter 5.

We assessed the performance of 4 different two-class classifiers: a support vector machine with a quadratic and a linear kernel (SVM2, SVM), C4.5 decision tree (DT), and Naïve Bayes (NB). To evaluate each classifier and the different approaches, we performed 9 leave-one-block-out cross-validations for each subject (see figures in *Materials and methods* section). For each cross-validation the training data contained 8 out of 9 task-rest blocks and the testing data the remaining block. For the connectivity-based approaches, FC during the rest and movies conditions were estimated from the training data only.

For all classifiers and approaches, we assessed whether each scan in the held-out test data was correctly classified (1 =correct, 0 =wrong), and concatenated the results across all folds.

We compared the performances of two classifiers across all subjects using McNemar's test (paired binary outcomes), and report the percentage of correctly classified scans across folds and subjects.

Additional discussion

Using simulated data showed that connectivity-based approaches were most beneficial for the deci-

sion tree and Naïve Bayes classifiers (see *Results - simulations* section). Linear and quadratic SVMs performed best overall, however, and gained less from the connectivity-based approach. As a side note we remark that even small differences in accuracy were significant because of the large number of samples: We compared the performances using scan-wise outcomes concatenated across subjects and thus small differences became significant.

Using experimental fMRI data we observed similar accuracy across all approaches and classifiers (see *Results - fMRI dataset* section). For each classifier, it is possible to compute a weight map representing which features (i.e., regions) have most discriminative power in separating the two cognitive states. Such a mapping for an example subject subject shows that with the activity-based approach information is concentrated in visual (calcarine, fusiform), emotional (amygdala) and prefrontal regions (ACC, orbitofrontal). The wavelet approach indicates scale-dependent importance of brain regions: For example, higher visual regions (fusiform, inferior and superior occipital cortex) are more important at coarse scales and the primary visual cortex (calcarine, lingual gyrus) at fine scales.

Conclusion

Incorporating connectivity information did not boost classification performance. While the multi-scale mapping of discriminative regions can provide insights not accessible with an activity-based approach, it also requires a much more complex analysis pipeline.

FMRI decoding of cognitive states using functional connectivity

Nora Leonardi[1,2] [nora.leonardi@epfl.ch], **Jonas Richiardi**[1,2], **Dimitri Van De Ville**[1,2]

1: Medical Image Processing Laboratory, Institute of Bioengineering, Ecole Polytechnique Fédérale de Lausanne (EPFL)
2: Medical Image Processing Laboratory, Department of Radiology and Medical Informatics, University of Geneva

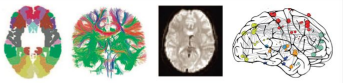
Key points

- State-of-the-art classification of instantaneous cognitive states relies solely on brain activity
- We propose to couple brain activity with state-specific functional connectivity to continuously decode a subject's cognitive state
- Simulations showed that the integration of connectivity information enhances the decoding accuracy
- Application to fMRI data (movies and rest) showed an enhanced accuracy for only a few subjects
- Proposed decoding method allows novel insights into the scale at which a brain region's activity is discriminative between cognitive states, i.e. localized versus distributed activity

Background

Functional connectivity as a signature of cognitive states

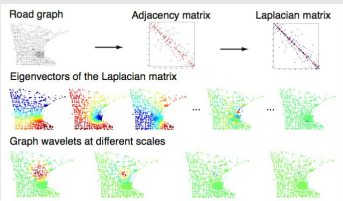
- The human brain is a **complex network** with highly structured connectivity and multiscale organization
- Patterns of functional connectivity can be used to decode a subject's cognitive state [1]



The constituents of brain graphs: anatomically defined nodes and structural (using diffusion imaging) or functional connections (using e.g. functional magnetic resonance imaging); example of a brain graph (from left to right) [2]

Graph wavelets adapted to a cognitive state's functional connectivity

- Wavelets provide a **multiscale** view of the data, where smooth (e.g. background) and sharp features (e.g. transient events) are represented at different scales.
- Graph wavelets and the graph Fourier transform are **adapted** to the underlying graph



A graph can be represented by its adjacency and Laplacian matrix (top), the eigenvectors of the Laplacian matrix are a graph structure-adapted basis (middle), graph wavelets diffuse along the graph structure (bottom, coarse to fine scale)

- Applying graph wavelets to fMRI data will capture interactions between the dynamics and structure of brain activity [3,4]

Objective

Capitalize on the **spatiotemporal structure of brain activity** to decode a subject's cognitive states: Can connectivity information help to decode cognitive states?

Materials and methods

Simulations

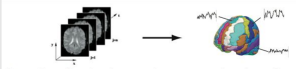
- 3 groups of simulations with connectivity and activity changes between 2 conditions

Application to fMRI data

- 15 healthy subjects scanned at 3 Tesla [5]
- Block-design with 2 conditions: movies, rest

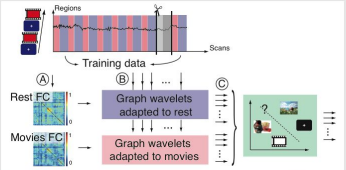
Data analysis

- Extraction of 90 regional mean time courses



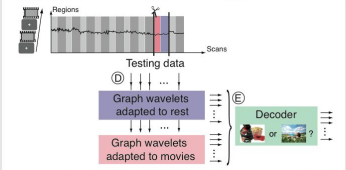
Extraction of regional mean time courses based on the automatic anatomical labeling (AAL) atlas

- **4 classification methods tested:** quadratic and linear SVM (SVM2, SVM), decision tree (DT) and Naïve Bayes (NB)
- **Leave-one-block out cross-validation**
  - **Train** the 'decoder' on the training data



Train the 'decoder': A) Estimate functional connectivity (FC) of the two cognitive states, B) Decompose the fMRI data using graph wavelets constructed from rest and movies FC, C) Train the 'decoder': Given the brain activity at a single scan, what is the subject's cognitive state?

  - **Evaluate** on novel (testing) data

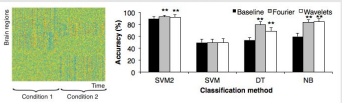


Evaluation of the 'decoder' on novel data: D) Decompose the fMRI data (using FC estimated in training data), E) Classify each scan and evaluate the performance

- **Compare** classification performance of **wavelets** and **Fourier** to a **baseline**, i.e. classification using the regional mean time courses (leaving out steps A, B and D)

Results - simulations

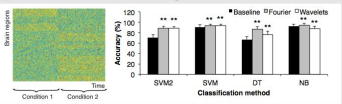
1) Differences in connectivity



Example of simulation (left), decoding accuracy; error bars represent the standard deviation, * a statistically significant difference to the baseline at the 5% level and ** at the 1% level (right)

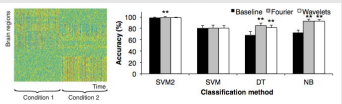
- Only SVM2 discriminated the 2 conditions
- Including connectivity improved the accuracy of quadratic SVM, NB and DT

2) Differences in activity levels



- SVM and NB discriminated well
- Including connectivity improved the accuracy of all classifiers (except for NB + wavelets)

3) Differences in activity and connectivity



- SVMs, especially SVM2, discriminated well
- Including connectivity improved the accuracy of SVM2, DCT and NB

Results - fMRI dataset

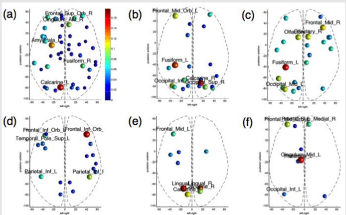
Decoding accuracy

- Accuracy was high for all classifiers, but highly variable across subjects as previously observed [6]
- Integrating connectivity information improved the decoding accuracy for a few subjects

Accuracy (%)	SVM2	SVM	DT	NB
Baseline	93.8	93.4	86	92.4
Fourier	94.4** (+2)	93.1** (+2/-2)	85.7 (+1/-2)	92.7* (-4)
Wavelets	94.6** (+3)	93.8 (+2/-1)	89.7** (+7)	93** (+2)

Average accuracy of the 4 classifiers across all 15 subjects and in brackets the number of subjects with significantly improved or worsened accuracy at the 5% level (+/- respectively)

Mapping discriminative regions at multiple scales



Axial view of example MNI coordinate maps for subject 13 depicting the relative discriminative power of each region in (a) the baseline and (b-f) wavelets NB classifier. (b-f) show the power in one of the 5 scales (from coarse to fine) of the graph wavelet decomposition

References

1. J. Richiardi, H. Eryilmaz, S. Schwartz, P. Vuilleumier, D. Van De Ville, Decoding brain states from fMRI connectivity graphs. *Neuroimage* 2011, 56(2):616-626
2. D.S. Bassett, M.S. Gazzaniga, Understanding complexity in the human brain. *Trends Cogn Sci* 2011, 15(5):200-209
3. D. Hammond, P. Vandergheynst, R. Gribonval, Wavelets on graphs via spectral graph theory. *Appl Comp Harm Anal* 2011, 30:129-150
4. N. Leonardi, D. Van De Ville, Wavelet frames on graphs defined by fMRI functional connectivity. In *ISBI* 2011, 2136-2139
5. H. Eryilmaz, D. Van De Ville, S. Schwartz, P. Vuilleumier, Impact of transient emotions on functional connectivity during subsequent resting state: a wavelet correlation approach. *Neuroimage* 2011, 54(3):2481-2491
6. T.M. Mitchell, R. Hutchinson, R.S. Niculescu, F. Pereira, X. Wang, M. Just, and S. Newman, Learning to decode cognitive states from brain images. *Machine Learning* 2004 57:145-175

Acknowledgements

This work was supported in part by the Swiss National Science Foundation (grant PP00P2-123438), in part by the Société Académique de Genève and the FOREMANE foundation, and in part by the Center for Biomedical Imaging (CIBM) of the Geneva and Lausanne Universities, EPFL, and the Leenaards and Louis-Jeantet foundations.

E Supplementary material for Chapter 6

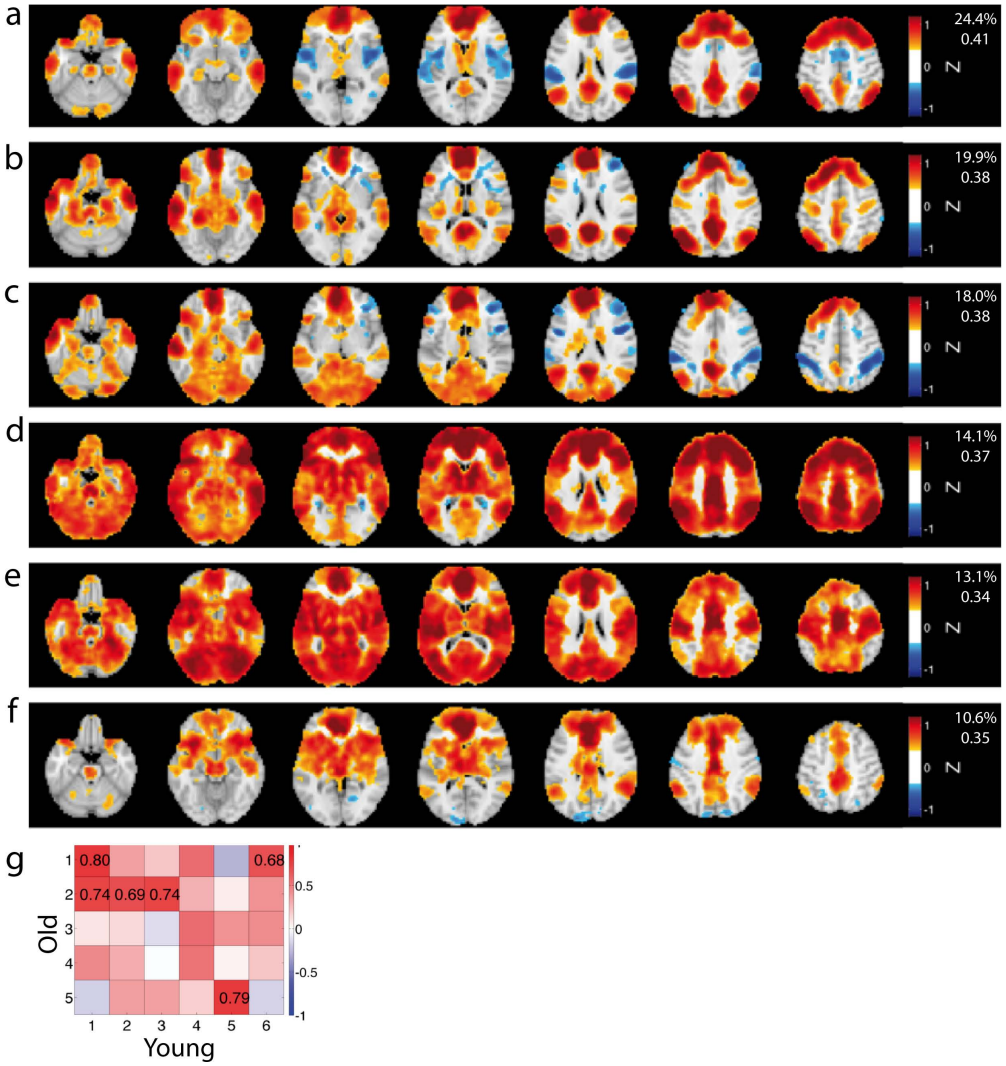


Figure E.1 – (a-f) 6 reproducible CAPs were identified in young subjects. (g) Spatial correlation coefficient between the 5 CAPs estimated in old subjects and the 6 CAPs estimated in young subjects. Pearson r value is indicated if above the reproducibility threshold in older subjects. The 2 DMN-related CAPs (CAPs 1 and 2 in old subjects) and visual-motion-CAP (CAP 5) appear in both groups, while the prefrontal and superior-DMN CAPs show low correlation. 4 of the 6 CAPs in young subjects correlate with the 2 DMN-related CAPs of old subjects. These findings confirm the differences observed in Fig. 6.8.

Bibliography

Achard S, Salvador R, Whitcher B, Suckling J and Bullmore E. (2006). A resilient, low-frequency, small-world human brain functional network with highly connected association cortical hubs. *J Neurosci*, **26**(1), 63–72.

Achard S and Bullmore E. (2007). Efficiency and cost of economical brain functional networks. *PLoS Comput Biol*, **3**(2), e17.

Afshin-Pour B, Hossein-Zadeh GA, Strother SC and Soltanian-Zadeh H. (2012). Enhancing reproducibility of fMRI statistical maps using generalized canonical correlation analysis in NPAIRS framework. *NeuroImage*, **60**(4), 1970–1981.

Aharon M, Elad M and Bruckstein A. (2006). K -SVD: An algorithm for designing overcomplete dictionaries for sparse representations. *Signal Processing, IEEE Transactions on*, **54**(11), 4311–4322.

Ahn YY, Bagrow JP and Lehmann S. (2010). Link communities reveal multiscale complexity in networks. *Nature*, **466**(7307), 761–764.

Albert NB, Robertson EM and Miall RC. (2009). The resting human brain and motor learning. *Curr Biol*, **19**(12), 1023–1027.

Allen EA, Eichele T, Wu L and Calhoun VD. (2013). EEG signatures of functional connectivity states. In *Proc. Organization for Human Brain Mapping (OHBM)*, 1, Beijing, China.

Allen EA, Damaraju E, Plis SM, Erhardt EB, Eichele T and Calhoun VD. (2014). Tracking whole-brain connectivity dynamics in the resting state. *Cereb Cortex*, **24**(3), 663–676.

Andrews-Hanna JR, Reidler JS, Sepulcre J, Poulin R and Buckner RL. (2010). Functional-anatomic fractionation of the brain's default network. *Neuron*, **65**(4), 550–562.

Andrews-Hanna JR, Snyder AZ, Vincent JL, Lustig C, Head D, Raichle ME and Buckner RL. (2007). Disruption of large-scale brain systems in advanced aging. *Neuron*, **56**(5), 924–935.

Appellof CJ and Davidson ER. (1981). Strategies for analyzing data from video fluorometric monitoring of liquid chromatographic effluents. *Anal Chem*, **53**(13), 2053–2056.

Arieli A, Sterkin A, Grinvald A and Aertsen A. (1996). Dynamics of ongoing activity: explanation of the large variability in evoked cortical responses. *Science*, **273**(5283), 1868–1871.

Ashburner J and Friston KJ. (2005). Unified segmentation. *NeuroImage*, **26**(3), 839–851.

Auscher P. Wavelet bases for l2 r with rational dilation factor. In *Wavelets and their applications*, 439–451. Jones and Bartlett, Boston, MA. (1992).

Bandettini PA, Wong EC, Hinks RS, Tikofsky RS and Hyde JS. (1992). Time course epi of human brain function during task activation. *Magn Reson Med*, **25**(2), 390–397.

Barkhof F. (2002). The clinico-radiological paradox in multiple sclerosis revisited. *Curr Opin Neurol*, **15** (3), 239–245.

Bassett DS and Gazzaniga MS. (2011). Understanding complexity in the human brain. *Trends Cogn Sci*, **15**(5), 200–209.

Bassett DS, Wymbs NF, Porter MA, Mucha PJ and Grafton ST. (2014). Cross-linked structure of network evolution. *Chaos*, **24**, 013112.

Bassett D, Wymbs N, Porter M, Mucha P, Carlson J and Grafton S. (2011). Dynamic reconfiguration of human brain networks during learning. *Proc Natl Acad Sci U S A*, **108**(18), 7641—-7646.

Baumgartner R, Scarth G, Teichtmeister C, Somorjai R and Moser E. (1997). Fuzzy clustering of gradient-echo functional mri in the human visual cortex. part i: reproducibility. *J Magn Reson Imaging*, **7**(6), 1094–1101.

Baune A, Sommer FT, Erb M, Wildgruber D, Kardatzki B, Palm G and Grodd W. (1999). Dynamical cluster analysis of cortical fmri activation. *NeuroImage*, **9**(5), 477–489.

Baussard A, Nicolier F and Truchetet F. (2004). Rational multiresolution analysis and fast wavelet transform: application to wavelet shrinkage denoising. *Signal Process.*, **84**(10), 1735–1747.

Beckmann CF, Mackay CE, Filippini N and Smith SM. (2009). Group comparison of resting-state FMRI data using multi-subject ICA and dual regression. In *Proc. Organization for Human Brain Mapping (OHBM)*, 1, San Francisco, CA.

Beckmann CF and Smith S. (2004). Tensorial extensions of independent component analysis for multisubject FMRI analysis. *NeuroImage*, **25**, 294–311.

Beckmann CF, De Luca M, Devlin JT and Smith SM. (2005). Investigations into resting-state connectivity using independent component analysis. *Philos Trans R Soc Lond B Biol Sci*, **360**(1457), 1001–1013.

Behjat H, Leonardi N and Van De Ville D. (2013). Statistical parametric mapping of functional MRI data using wavelets adapted to the cerebral cortex. In *Proc IEEE International Symposium on Biomedical Imaging (ISBI)*, 1058–1061, San Francisco, CA.

Benedetto J and Fickus M. (2003). Finite normalized tight frames. *Adv Comp Math*, **18**(2), 357–385.

Benedict RH, Schretlen D, Groninger L and Brandt J. (1998). Hopkins verbal learning test-revised: Normative data and analysis of inter-form and test-retest reliability. *The Clinical Neuropsychologist*, **12**(1), 43–55.

Biswal B, Yetkin FZ, Haughton VM and Hyde JS. (1995). Functional connectivity in the motor cortex of resting human brain using echo-planar mri. *Magn Reson Med*, **34**(4), 537–541.

Bjornsson H and Venegas SA. *A Manual for EOF and SVD analyses of Climatic Data*, 1997.

Boly M, Balteau E, Schnakers C, Degueldre C, Moonen G, Luxen A, Phillips C, Peigneux P, Maquet P and Laureys S. (2007). Baseline brain activity fluctuations predict somatosensory perception in humans. *Proc Natl Acad Sci U S A*, **104**(29), 12187–12192.

Boly M, Phillips C, Tshibanda L, Vanhaudenhuyse A, Schabus M, Dang-Vu TT, Moonen G, Hustinx R, Maquet P and Laureys S. (2008). Intrinsic brain activity in altered states of consciousness: how conscious is the default mode of brain function? *Ann N Y Acad Sci*, **1129**, 119–129.

Bonavita S, Gallo A, Sacco R, Corte M, Bisecco A, Docimo R, Lavorgna L, Corbo D, Di Costanzo A, Tortora F, Cirillo M, Esposito F and Tedeschi G. (2011). Distributed changes in default-mode resting-state connectivity in multiple sclerosis. *Mult Scler*, **17**(4), 411–422.

Brandt J. (1991). The hopkins verbal learning test: Development of a new memory test with six equivalent forms. *Clinical Neuropsychologist*, **5**(2), 125–142.

Breakspear M, Bullmore ET, Aquino K, Das P and Williams LM. (2006). The multiscale character of evoked cortical activity. *NeuroImage*, **30**(4), 1230–1242.

Britz J, Van De Ville D and Michel CM. (2010). Bold correlates of eeg topography reveal rapid resting-state network dynamics. *NeuroImage*, **52**(4), 1162–1170.

Bro R. (1997). PARAFAC. Tutorial and applications. *Chem Intell Lab Syst*, **38**, 149–171.

Buckner RL, Andrews-Hanna JR and Schacter DL. (2008). The brain's default network: anatomy, function, and relevance to disease. *Ann N Y Acad Sci*, **1124**, 1–38.

Buckner RL, Krienen FM and Yeo BTT. (2013). Opportunities and limitations of intrinsic functional connectivity mri. *Nat Neurosci*, **16**(7), 832–837.

Buckner R, Sepulcre J, Talukdar T, Krienen F, Liu H, Hedden T, Andrews-Hanna J, Sperling R and Johnson K. (2009). Cortical hubs revealed by intrinsic functional connectivity: mapping, assessment of stability, and relation to alzheimer's diseases. *J Neurosci*, **29**(6), 1860–1873.

Buldyrev S, Parshani R, Paul G, Stanley H and Havlin S. (2010). Catastrophic cascade of failures in interdependent networks. *Nature*, **464**(7291), 1025–1028.

Bullmore E, Fadili J, Maxim V, Sendur L, Whitcher B, Suckling J, Brammer M and Breakspear M. (2004). Wavelets and functional magnetic resonance imaging of the human brain. *NeuroImage*, **23**, S234–S249.

Bullmore E and Sporns O. (2009). Complex brain networks: graph theoretical analysis of structural and functional systems. *Nat Rev Neurosci*, **10**(3), 186–198.

Cabeza R and St Jacques P. (2007). Functional neuroimaging of autobiographical memory. *Trends Cogn Sci*, **11**(5), 219–227.

Calhoun V and Adali T. (2012). Multi-subject independent component analysis of fMRI: A decade of intrinsic networks, default mode, and neurodiagnostic discovery. *IEEE Rev Biomed Eng*, **5**(99), 60–73.

Calhoun VD, Eichele T, Adalı T and Allen EA. (2012). Decomposing the brain: components and modes, networks and nodes. *Trends Cogn Sci*, **16**(5), 255–256.

Carroll J and Chang J. (1970). Analysis of individual differences in multidimensional scaling via an n-way generalization of "eckart-young" decomposition. *Psychometrika*, **35**, 283–319.

Cattaneo C. (1997). The spectrum of the continuous Laplacian on a graphs. *Monatshefte fuer Mathematik*, **124**, 215–235.

Chang C and Glover G. (2010). Time-frequency dynamics of resting-state brain connectivity measured with fMRI. *NeuroImage*, **50**(1), 81–98.

Chang C and Glover GH. (2009). Effects of model-based physiological noise correction on default mode network anti-correlations and correlations. *NeuroImage*, **47**(4), 1448–1459.

Chang C, Liu Z, Chen MC, Liu X and Duyn JH. (2013a). EEG correlates of time-varying BOLD functional connectivity. *NeuroImage*, **72**, 227–236.

Chang C, Metzger CD, Glover GH, Duyn JH, Heinze HJ and Walter M. (2013b). Association between heart rate variability and fluctuations in resting-state functional connectivity. *NeuroImage*, **68**, 93 – 104.

Chao-Gan Y and Yu-Feng Z. (2010). DPARSF: A MATLAB toolbox for “pipeline" data analysis of resting-state fMRI. *Front Syst Neurosci*, **4**, 13.

Chung F. (1997). *Spectral graph theory*. AMS, Providence, RI.

Coifman R and Maggioni M. (2006). Diffusion wavelets. *Appl Comput Harmon Anal*, **21**(21), 53–94.

Cole DM, Smith SM and Beckmann CF. (2010). Advances and pitfalls in the analysis and interpretation of resting-state fmri data. *Front Syst Neurosci*, **4**, 8.

Compston A and Coles A. (2008). Multiple sclerosis. *Lancet*, **372**(9648), 1502–1517.

Cordes D, Haughton VM, Arfanakis K, Wendt GJ, Turski PA, Moritz CH, Quigley MA and Meyerand ME. (2000). Mapping functionally related regions of brain with functional connectivity mr imaging. *AJNR Am J Neuroradiol*, **21**(9), 1636–1644.

Cordes D and Nandy RR. (2006). Estimation of the intrinsic dimensionality of fmri data. *NeuroImage*, **29**(1), 145 – 154.

Craddock RC, Holtzheimer 3rdPE, Hu XP and Mayberg HS. (2009). Disease state prediction from resting state functional connectivity. *Magn Reson Med*, **62**(6), 1619–1628.

Craddock RC, James GA, Holtzheimer 3rdPE, Hu XP and Mayberg HS. (2012). A whole brain fmri atlas generated via spatially constrained spectral clustering. *Hum Brain Mapp*, **33**(8), 1914–1928.

Cribben I, Haraldsdottir R, Atlas LY, Wager TD and Lindquist MA. (2012). Dynamic connectivity regression: Determining state-related changes in brain connectivity. *NeuroImage*, **61**, 907–920.

Damoiseaux JS, Beckmann CF, Arigita EJS, Barkhof F, Scheltens P, Stam CJ, Smith SM and Rombouts SARB. (2008). Reduced resting-state brain activity in the "default network" in normal aging. *Cereb Cortex*, **18**(8), 1856–1864.

Damoiseaux JS, Rombouts SARB, Barkhof F, Scheltens P, Stam CJ, Smith SM and Beckmann CF. (2006). Consistent resting-state networks across healthy subjects. *Proc Natl Acad Sci U S A*, **103**(37), 13848–13853.

de Chastelaine M, Wang TH, Minton B, Muftuler LT and Rugg MD. (2011). The effects of age, memory performance, and callosal integrity on the neural correlates of successful associative encoding. *Cereb Cortex*, **21**(9), 2166–2176.

De Domenico M, Solé-Ribalta A, Cozzo E, Kivelä M, Moreno Y, Porter MA, Gómez S and Arenas A. (2013). Mathematical formulation of multilayer networks. *Phys Rev X*, **3**, 041022.

De Lathauwer L, De Moor B and Vandewalle J. (2000a). A multilinear singular value decomposition. *SIAM J Matrix Anal Appl*, **21**, 1253–1278.

De Lathauwer L, De Moor B and Vandewalle J. (2000b). On the best rank-1 and rank-($R_1, R_2, ..., R_N$) approximation of higher-order tensors. *SIAM J Matrix Anal Appl*, **21**(4), 1324–1342.

De Luca M, Smith S, De Stefano N, Federico A and Matthews PM. (2005). Blood oxygenation level dependent contrast resting state networks are relevant to functional activity in the neocortical sensorimotor system. *Exp Brain Res*, **167**(4), 587–594.

Deco G, Jirsa VK and McIntosh AR. (2013). Resting brains never rest: computational insights into potential cognitive architectures. *Trends Neurosci*, **36**(5), 268–274.

Di Zenzo S. (1986). A note on the gradient of a multi-image. *Comput Vision Graph Image Process*, **33**(1), 116–125.

Diekelmann S and Born J. (2010). The memory function of sleep. *Nat Rev Neurosci*, **11**(2), 114–126.

Dogonowski AM, Siebner HR, Sørensen PS, Wu X, Biswal B, Paulson OB, Dyrby TB, Skimminge A, Blinkenberg M and Madsen KH. (2013). Expanded functional coupling of subcortical nuclei with the motor resting-state network in multiple sclerosis. *Mult Scler*, **19**, 559–566.

Dong X, Frossard P, Vandergheynst P and Nefedov N. (2012). Clustering with multi-layer graphs: A spectral perspective. *IEEE Trans Signal Process*, **60**, 5820–5831.

Dosenbach NUF, Nardos B, Cohen AL, Fair DA, Power JD, Church JA, Nelson SM, Wig GS, Vogel AC, Lessov-Schlaggar CN, Barnes KA, Dubis JW, Feczko E, Coalson RS, Pruett JrJR, Barch DM, Petersen SE and Schlaggar BL. (2010). Prediction of individual brain maturity using fMRI. *Science*, **329**(5997), 1358–1361.

Doucet G, Naveau M, Petit L, Delcroix N, Zago L, Crivello F, Jobard G, Tzourio-Mazoyer N, Mazoyer B, Mellet E and Joliot M. (2011). Brain activity at rest: a multiscale hierarchical functional organization. *J Neurophysiol*, **105**(6), 2753–2763.

Dronkers NF, Plaisant O, Iba-Zizen MT and Cabanis EA. (2007). Paul broca's historic cases: high resolution mr imaging of the brains of leborgne and lelong. *Brain*, **130**(Pt 5), 1432–1441.

Dunlavy D, Kolda T and Kegelmeyer W. (2011). *Multilinear Algebra for Analyzing Data with Multiple Linkages*, chapter 7, 85–114. SIAM.

Eavani H, Filipovych R, Davatzikos C, Satterthwaite TD, Gur RE and Gur RC. (2012). Sparse dictionary learning of resting state fmri networks. In *Proc IEEE Pattern Recognition in NeuroImaging (PRNI)*, PRNI '12, 73–76, Washington, DC, USA. IEEE Computer Society.

Eavani H, Satterthwaite T, Gur R, Gur R and Davatzikos C. Unsupervised learning of functional network dynamics in resting state fMRI. In Gee J, Joshi S, Pohl K, Wells W and ZÃ¶llei L, editors, *Information Processing in Medical Imaging*, volume 7917 of *Lecture Notes in Computer Science*, 426–437. Springer Berlin Heidelberg. (2013). ISBN 978-3-642-38867-5.

Eichele T, Debener S, Calhoun VD, Specht K, Engel AK, Hugdahl K, von Cramon DY and Ullsperger M. (2008). Prediction of human errors by maladaptive changes in event-related brain networks. *Proc Natl Acad Sci U S A*, **105**(16), 6173–6178.

Ekman M, Derrfuss J, Tittgemeyer M and Fiebach CJ. (2012). Predicting errors from reconfiguration patterns in human brain networks. *Proc Natl Acad Sci U S A*, **109**(41), 16714–16719.

Eldén L. (2007). *Matrix methods in data mining and pattern recognition*. SIAM, Philadelphia, PA.

Eryilmaz H, Van De Ville D, Schwartz S and Vuilleumier P. (2011). Impact of transient emotions on functional connectivity during subsequent resting state: A wavelet correlation approach. *NeuroImage*, **54**(3), 2481–2491.

Euston DR, Gruber AJ and McNaughton BL. (2012). The role of medial prefrontal cortex in memory and decision making. *Neuron*, **76**(6), 1057–1070.

Ferrarini L, Veer IM, Baerends E, van Tol MJ, Renken RJ, van der Wee NJA, Veltman DJ, Aleman A, Zitman FG, Penninx BWJH, van Buchem MA, Reiber JHC, Rombouts SARB and Milles J. (2009). Hierarchical functional modularity in the resting-state human brain. *Hum Brain Mapp*, **30**(7), 2220–2231.

Ferreira LK and Busatto GF. (2013). Resting-state functional connectivity in normal brain aging. *Neurosci Biobehav Rev*, **37**(3), 384–400.

Filippi M and Agosta F. (2010). Imaging biomarkers in multiple sclerosis. *J Magn Reson Imaging*, **31**(4), 770–788.

Fox MD and Greicius MD. (2010). Clinical applications of resting state functional connectivity. *Front Syst Neurosci*, **4**, 19.

Fox MD and Raichle ME. (2007). Spontaneous fluctuations in brain activity observed with functional magnetic resonance imaging. *Nat Rev Neurosci*, **8**(9), 700–711.

Fox MD, Snyder AZ, Zacks JM and Raichle ME. (2006). Coherent spontaneous activity accounts for trial-to-trial variability in human evoked brain responses. *Nat Neurosci*, **9**(1), 23–25.

Friedman J and Tillich JP. (2004). Wave equations for graphs and the edge-based Laplacian. *Pacific J. Math.*, **216**, 229–266.

Friston KJ, Frith CD, Liddle PF and Frackowiak RS. (1993). Functional connectivity: the principal-component analysis of large (pet) data sets. *J Cereb Blood Flow Metab*, **13**(1), 5–14.

Friston KJ, Holmes AP, Poline JB, Grasby PJ, Williams SC, Frackowiak RS and Turner R. (1995). Analysis of fmri time-series revisited. *NeuroImage*, **2**(1), 45–53.

Friston KJ, Jezzard P, Frackowiak RSJ and Turner R. (1994). *Functional MRI of the Brain*, chapter Characterizing focal and distributed physiological changes with MRI and PET, 207–216. Society of Magnetic Resonance in Medicine, Berkeley, CA.

Fu L, Matthews PM, De Stefano N, Worsley KJ, Narayanan S, Francis GS, Antel JP, Wolfson C and Arnold DL. (1998). Imaging axonal damage of normal-appearing white matter in multiple sclerosis. *Brain*, **121**, 103–113.

Garrett DD, Kovacevic N, McIntosh AR and Grady CL. (2010). Blood oxygen level-dependent signal variability is more than just noise. *J Neurosci*, **30**(14), 4914–4921.

Giorgio A, Santelli L, Tomassini V, Bosnell R, Smith S, De Stefano N and Johansen-Berg H. (2010). Age-related changes in grey and white matter structure throughout adulthood. *Neuroimage*, **51**(3), 943–951.

Glover GH, Li TQ and Ress D. (2000). Image-based method for retrospective correction of physiological motion effects in fmri: Retroicor. *Magn Reson Med*, **44**(1), 162–167.

Gómez S, Díaz-Guilera A, Gómez-Gardeñes J, Pérez-Vicente CJ, Moreno Y and Arenas A. (2013). Diffusion dynamics on multiplex networks. *Phys Rev Lett*, **110**, 028701.

Gonzalez-Castillo J, Hoy C, Handwerker DA, Robinson M and Band . (2013). Detection of consistent cognitive processing at the single subject level using whole-brain functional connectivity. In *Society for Neuroscience*, San Diego, CA.

Good CD, Johnsrude IS, Ashburner J, Henson RN, Friston KJ and Frackowiak RS. (2001). A voxel-based morphometric study of ageing in 465 normal adult human brains. *NeuroImage*, **14**(1 Pt 1), 21–36.

Grady CL, Maisog JM, Horwitz B, Ungerleider LG, Mentis MJ, Salerno JA, Pietrini P, Wagner E and Haxby JV. (1994). Age-related changes in cortical blood flow activation during visual processing of faces and location. *J Neurosci*, **14**(3 Pt 2), 1450–1462.

Grady C. (2012). The cognitive neuroscience of ageing. *Nat Rev Neurosci*, **13**(7), 491–505.

Greicius MD. (2008). Resting-state functional connectivity in neuropsychiatric disorders. *Curr Opin Neurol*, **21**(4), 424–430.

Greicius MD, Kiviniemi V, Tervonen O, Vainionpää V, Alahuhta S, Reiss AL and Menon V. (2008). Persistent default-mode network connectivity during light sedation. *Hum Brain Mapp*, **29**(7), 839–847.

Greicius MD, Krasnow B, Reiss AL and Menon V. (2003). Functional connectivity in the resting brain: a network analysis of the default mode hypothesis. *Proc Natl Acad Sci U S A*, **100**(1), 253–258.

Greicius MD, Srivastava G, Reiss AL and Menon V. (2004). Default-mode network activity distinguishes alzheimer's disease from healthy aging: Evidence from functional mri. *Proc Natl Acad Sci U S A*, **101** (13), 4637–4642.

Greicius MD, Supekar K, Menon V and Dougherty RF. (2009). Resting-state functional connectivity reflects structural connectivity in the default mode network. *Cereb Cortex*, **19**(1), 72–78.

Gunning-Dixon FM, Brickman AM, Cheng JC and Alexopoulos GS. (2009). Aging of cerebral white matter: a review of mri findings. *Int J Geriatr Psychiatry*, **24**(2), 109–117.

Hackmack K, Paul F, Weygandt M, Allefeld C, Haynes JD and Initiative ADN. (2012). Multi-scale classification of disease using structural MRI and wavelet transform. *NeuroImage*, **62**(1), 48–58.

Hammond D, Raoaroor K, Jacques L and Vandergheynst P. (2010). Image denoising with nonlocal spectral graph wavelets. In *SIAM Conference on Imaging Science*.

Hammond D, Vandergheynst P and Gribonval R. (2011). Wavelets on graphs via spectral graph theory. *Appl Comput Harmon Anal*, **30**, 129–150.

Hampson M, Peterson BS, Skudlarski P, Gatenby JC and Gore JC. (2002). Detection of functional connectivity using temporal correlations in mr images. *Hum Brain Mapp*, **15**(4), 247–262.

Handwerker DA, Roopchansingh V, Gonzalez-Castillo J and Bandettini PA. (2012). Periodic changes in fMRI connectivity. *NeuroImage*, **63**(3), 1712–1719.

Harmelech T and Malach R. (2013). Neurocognitive biases and the patterns of spontaneous correlations in the human cortex. *Trends Cogn Sci*, **17**(12), 606–615.

Harshman R. (1970). Foundations of the PARAFAC procedure: Models and conditions for an "explanatory" multi-modal factor analysis. *UCLA Working Papers in Phonetics*, **16**(1), 1–84.

Harshman R and Hong S. (2002). 'Stretch' vs 'slice' methods for representing three-way structure via matrix notation. *J Chemometrics*, **16**(4), 198–205.

Hawellek DJ, Hipp JF, Lewis CM, Corbetta M and Engel AK. (2011). Increased functional connectivity indicates the severity of cognitive impairment in multiple sclerosis. *Proc Natl Acad Sci U S A*, **108** (47), 19066–19071.

Haxby JV, Gobbini MI, Furey ML, Ishai A, Schouten JL and Pietrini P. (2001). Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science*, **293**(5539), 2425–2430.

He BJ, Snyder AZ, Zempel JM, Smyth MD and Raichle ME. (2008). Electrophysiological correlates of the brain's intrinsic large-scale functional architecture. *Proc Natl Acad Sci U S A*, **105**(41), 16039–16044.

He J, Carmichael O, Fletcher E, Singh B, Iosif AM, Martinez O, Reed B, Yonelinas A and Decarli C. (2012). Influence of functional connectivity and structural mri measures on episodic memory. *Neurobiol Aging*, **33**(11), 2612–2620.

Held S, Storath M, Massopust P and Forster B. (2010). Steerable wavelet frames based on the riesz transform. *IEEE Trans Image Process*, **19**, 653 –667.

Hesselmann G, Kell CA, Eger E and Kleinschmidt A. (2008). Spontaneous local variations in ongoing neural activity bias perceptual decisions. *Proc Natl Acad Sci U S A*, **105**(31), 10984–10989.

Honey CJ, Kötter R, Breakspear M and Sporns O. (2007). Network structure of cerebral cortex shapes functional connectivity on multiple time scales. *Proc Natl Acad Sci U S A*, **104**(24), 10240–10245.

Hotelling H. (1936). Relations between two sets of variates. *Biometrika*, **28**, 321–377.

Hu C, Cheng L, Sepulcre J, El Fakhri G, Lu YM and Li Q. (2013). Matched signal detection on graphs: Theory and application to brain network classification. In *Proc International Conference on Information Processing in Medical Imaging (IPMI)*, 1–12, Asilomar, CA.

Huettel S, Song A and McCarthy G. (2004). *Functional magnetic resonance imaging*. Sinauer.

Hutchison RM, Womelsdorf T, Allen EA, Bandettini PA, Calhoun VD, Corbetta M, Della Penna S, Duyn JH, Glover GH, Gonzalez-Castillo J, Handwerker DA, Keilholz S, Kiviniemi V, Leopold DA, de Pasquale F, Sporns O, Walter M and Chang C. (2013a). Dynamic functional connectivity: promise, issues, and interpretations. *NeuroImage*, **80**, 360–378.

Hutchison RM, Womelsdorf T, Gati JS, Everling S and Menon RS. (2013b). Resting-state networks show dynamic functional connectivity in awake humans and anesthetized macaques. *Hum Brain Mapp*, **34**(9), 2154–77.

Jafri MJ, Pearlson GD, Stevens M and Calhoun VD. (2008). A method for functional network connectivity among spatially independent resting-state components in schizophrenia. *NeuroImage*, **39**(4), 1666–1681.

Jansen M, Nason GP and Silverman BW. (2009). Multiscale methods for data on graphs and irregular multidimensional situations. *J R Stat Soc Ser B Stat Methodol*, **71**(1), 97–125.

Jolliffe I. (1986). *Principal component analysis*. Springer.

Jones DT, Vemuri P, Murphy MC, Gunter JL, Senjem ML, Machulda MM, Przybelski SA, Gregg BE, Kantarci K, Knopman DS, Boeve BF, Petersen RC and Jack JrCR. (2012). Non-stationarity in the "resting brain's" modular architecture. *PLoS One*, **7**(6), e39731.

Kang J, Wang L, Yan C, Wang J, Liang X and He Y. (2011). Characterizing dynamic functional connectivity in the resting brain using variable parameter regression and Kalman filtering approaches. *NeuroImage*, **56**(3), 1222–1234.

Karahanoğlu FI and Van De Ville D. (2014). Investigating the spatial and temporal interactions in resting-state fMRI with Total Activation. In *Proc IEEE International Symposium on Biomedical Imaging (ISBI)*, 588–591, Beijing.

Karahanoğlu FI, Caballero-Gaudes C, Lazeyras F and Van de Ville D. (2013). Total activation: fmri deconvolution through spatio-temporal regularization. *NeuroImage*, **73**, 121–134.

Kettenring J. (1971). Canonical analysis of several sets of variables. *Biometrika*, **58**(3), 433–451.

Kim H. (2010). Dissociating the roles of the default-mode, dorsal, and ventral networks in episodic memory retrieval. *NeuroImage*, **50**(4), 1648–1657.

Kim WH, Singh V, Chung MK, Hinrichs C, Pachauri D, Okonkwo OC, Johnson SC and FtADNI. (2014). Multi-resolutional shape features via non-euclidean wavelets: Applications to statistical analysis of cortical thickness. *NeuroImage*.

Kitzbichler MG, Smith ML, Christensen SR and Bullmore E. (2009). Broadband criticality of human brain network synchronization. *PLoS Comput Biol*, **5**(3), e1000314.

Kivelä M, Arenas A, Barthelemy M, Gleeson JP, Moreno Y and Porter MA. (2013). Multilayer networks. *arXiv preprint*, arXiv:1309.7233.

Kiviniemi V, Kantola JH, Jauhiainen J, Hyvärinen A and Tervonen O. (2003). Independent component analysis of nondeterministic fMRI signal sources. *NeuroImage*, **19**(2 Pt 1), 253–260.

Kiviniemi V, Vire T, Remes J, Elseoud AA, Starck T, Tervonen O and Nikkinen J. (2011). A sliding time-window ICA reveals spatial variability of the default mode network in time. *Brain Connect*, **1**(4), 339–347.

Kolda TG and Bader BW. (2009). Tensor decompositions and applications. *SIAM Rev*, **51**(3), 455–500.

Krishnan A, Williams LJ, McIntosh AR and Abdi H. (2011). Partial Least Squares (PLS) methods for neuroimaging: a tutorial and review. *NeuroImage*, **56**(2), 455–475.

Kuhn HW. (1955). The hungarian method for the assignment problem. *Naval Research Logistics Quarterly*, **2**(1-2), 83–97.

Kurtzke JF. (1983). Rating neurologic impairment in multiple sclerosis: an expanded disability status scale (EDSS). *Neurology*, **33**(11), 1444–1452.

LaBar KS, Gitelman DR, Parrish TB and Mesulam M. (1999). Neuroanatomic overlap of working memory and spatial attention networks: a functional MRI comparison within subjects. *NeuroImage*, **10**(6), 695–704.

Laird AR, Fox PM, Eickhoff SB, Turner JA, Ray KL, McKay DR, Glahn DC, Beckmann CF, Smith SM and Fox PT. (2011). Behavioral interpretations of intrinsic connectivity networks. *J Cogn Neurosci*, **23**(12), 4022–4037.

Larson-Prior LJ, Zempel JM, Nolan TS, Prior FW, Snyder AZ and Raichle ME. (2009). Cortical network functional connectivity in the descent to sleep. *Proc Natl Acad Sci U S A*, **106**(11), 4489–4494.

Lee J, Jeong Y and Ye JC. (2013). Group sparse dictionary learning and inference for resting-state fmri analysis of alzheimer's disease. In *Proc IEEE International Symposium on Biomedical Imaging (ISBI)*, 540–543, San Francisco, CA.

Lee K, Tak S and Ye JC. (2011). A data-driven sparse glm for fmri analysis using sparse dictionary learning with mdl criterion. *Medical Imaging, IEEE Transactions on*, **30**(5), 1076–1089.

Leonardi N, Richiardi J, Gschwind M, Simioni S, Annoni JM, Schluep M, Vuilleumier P and Van De Ville D. (2013). Principal components of functional connectivity: a new approach to study dynamic brain connectivity during rest. *NeuroImage*, **83**, 937–950.

Leonardi N, Shirer WR, Greicius MD and Van de Ville D. (submitted). Disentangling dynamic networks by sparse combinations of functional connectivity patterns. *Hum Brain Mapp*.

Leonardi N and Van De Ville D. (2013a). Tight wavelet frames on multislice graphs. *IEEE Trans Signal Process*, **61**(13), 3357–3367.

Leonardi N and Van De Ville D. (2013b). Identifying network correlates of brain states using tensor decompositions of whole-brain dynamic functional connectivity. In *Proc IEEE International Workshop on Pattern Recognition in Neuroimaging (PRNI)*, 74–77, Philadelphia, PA.

Leonardi N and Van De Ville D. (2011). Wavelet frames on graphs defined by fMRI functional connectivity. In *Proc IEEE International Symposium on Biomedical Imaging (ISBI)*, 2136–2139, Chicago, IL.

Lewis CM, Baldassarre A, Committeri G, Romani GL and Corbetta M. (2009). Learning sculpts the spontaneous activity of the resting human brain. *Proc Natl Acad Sci U S A*, **106**(41), 17558–17563.

Li X, Zhu D, Jiang X, Jin C, Zhang X, Guo L, Zhang J, Hu X, Li L and Liu T. (2014). Dynamic functional connectomics signatures for characterization and differentiation of PTSD patients. *Hum Brain Mapp*, **35**(4), 1761–1778.

Liao JC, Boscolo R, Yang YL, Tran LM, Sabatti C and Roychowdhury VP. (2003). Network component analysis: reconstruction of regulatory signals in biological systems. *Proc Natl Acad Sci U S A*, **100**(26), 15522–15527.

Liu X, Chang C and Duyn JH. (2013). Decomposition of spontaneous brain activity into distinct fmri co-activation patterns. *Front Syst Neurosci*, **7**, 101.

Liu X and Duyn JH. (2013). Time-varying functional network information extracted from brief instances of spontaneous brain activity. *Proc Natl Acad Sci U S A*, **110**(11), 4392–4397.

Lowe MJ, Mock BJ and Sorenson JA. (1998). Functional connectivity in single and multislice echoplanar imaging using resting-state fluctuations. *NeuroImage*, **7**(2), 119–132.

Lowe MJ, Phillips MD, Lurito JT, Mattson D, Dzemidzic M and Mathews VP. (2002). Multiple sclerosis: low-frequency temporal blood oxygen level-dependent fluctuations indicate reduced functional connectivity initial results. *Radiology*, **224**(1), 184–192.

Lustig C, Snyder AZ, Bhakta M, O'Brien KC, McAvoy M, Raichle ME, Morris JC and Buckner RL. (2003). Functional deactivations: change with age and dementia of the alzheimer type. *Proc Natl Acad Sci U S A*, **100**(24), 14504–14509.

Mairal J, Bach F, Ponce J and Sapiro G. (2010). Online learning for matrix factorization and sparse coding. *J Mach Learn Res*, **11**, 19–60.

Majeed W, Magnuson M, Hasenkamp W, Schwarb H, Schumacher EH, Barsalou L and Keilholz SD. (2011). Spatiotemporal dynamics of low frequency BOLD fluctuations in rats and humans. *NeuroImage*, **54**(2), 1140–1150.

Makeig S, Debener S, Onton J and Delorme A. (2004). Mining event-related brain dynamics. *Trends Cogn Sci*, **8**(5), 204–210.

Mander BA, Rao V, Lu B, Saletin JM, Lindquist JR, Ancoli-Israel S, Jagust W and Walker MP. (2013). Prefrontal atrophy, disrupted nrem slow waves and impaired hippocampal-dependent memory in aging. *Nat Neurosci*, **16**(3), 357–364.

Marshall L, Helgadóttir H, Mölle M and Born J. (2006). Boosting slow oscillations during sleep potentiates memory. *Nature*, **444**(7119), 610–613.

Mazoyer B, Zago L, Mellet E, Bricogne S, Etard O, HoudÃ© O, Crivello F, Joliot M, Petit L and Tzourio-Mazoyer N. (2001). Cortical networks for working memory and executive functions sustain the conscious resting state in man. *Brain Research Bulletin*, **54**(3), 287 – 298.

McIntosh AR, Bookstein FL, Haxby JV and Grady CL. (1996). Spatial pattern analysis of functional brain images using partial least squares. *NeuroImage*, **3**, 143–157.

McIntosh AR, Kovacevic N and Itier RJ. (2008). Increased brain signal variability accompanies lower behavioral variability in development. *PLoS Comput Biol*, **4**(7), e1000106.

McKeown MJ, Makeig S, Brown GG, Jung TP, Kindermann SS, Bell AJ and Sejnowski TJ. (1998). Analysis of fMRI data by blind separation into independent spatial components. *Hum Brain Mapp*, **6**(3), 160–188.

Mehrkanoon S, Breakspear M and Boonstra T. (2014). Low-dimensional dynamics of resting-state cortical activity. *Brain Topogr*, **27**(3), 338–352.

Menon V. (2011). Large-scale brain networks and psychopathology: a unifying triple network model. *Trends Cogn Sci*, **15**(10), 483–506.

Meunier D, Achard S, Morcom A and Bullmore E. (2009a). Age-related changes in modular organization of human brain functional networks. *NeuroImage*, **44**(3), 715–723.

Meunier D, Lambiotte R, Fornito A, Ersche KD and Bullmore ET. (2009b). Hierarchical modularity in human brain functional networks. *Front Neuroinform*, **3**, 37.

Mevel K, Landeau B, Fouquet M, La Joie R, Villain N, Mézenge F, Perrotin A, Eustache F, Desgranges B and Chételat G. (2013). Age effect on the default mode network, inner thoughts, and cognitive abilities. *Neurobiol Aging*, **34**(4), 1292–1301.

Meyer Y. (1986). Principe d'incertitude, bases hilbertiennes et algèbres d'opérateurs. *Seminaire Bourbaki*, **662**, 209–223.

Minka TP. (2000). Automatic choice of dimensionality for PCA. In *Proc Neural Information Processing Systems (NIPS)*, 598–604.

Mucha PJ, Richardson T, Macon K, Porter MA and Onnela JP. (2010). Community structure in time-dependent, multiscale, and multiplex networks. *Science*, **328**(5980), 876–878.

Munkres J. (1957). Algorithms for the assignment and transportation problems. *Journal of the Society for Industrial and Applied Mathematics*, **5**(1), 32–38.

Narang SK and Ortega A. (2012). Perfect reconstruction two-channel wavelet filter banks for graph structured data. *IEEE Trans Signal Process*, **60**, 2786–2799.

Ng B, Abugharbieh R, Varoquaux G, Poline JB and Thirion B. (2011). Connectivity-informed fmri activation detection. *Med Image Comput Comput Assist Interv*, **14**(Pt 2), 285–292.

Nichols TE and Holmes AP. (2002). Nonparametric permutation tests for functional neuroimaging: a primer with examples. *Hum Brain Mapp*, **15**(1), 1–25.

Nir Y, Fisch L, Mukamel R, Gelbard-Sagiv H, Arieli A, Fried I and Malach R. (2007). Coupling between neuronal firing rate, gamma LFP, and BOLD fMRI is related to interneuronal correlations. *Curr Biol*, **17**(15), 1275–1285.

O'Sullivan M, Jones DK, Summers PE, Morris RG, Williams SC and Markus HS. (2001). Evidence for cortical "disconnection" as a mechanism of age-related cognitive decline. *Neurology*, **57**(4), 632–638.

Pantano P, Iannetti GD, Caramia F, Mainero C, Di Legge S, Bozzao L, Pozzilli C and Lenzi GL. (2002). Cortical motor reorganization after a single clinical attack of multiple sclerosis. *Brain*, **125**(Pt 7), 1607–1615.

Pascual-Marqui RD, Michel CM and Lehmann D. (1995). Segmentation of brain electrical activity into microstates: model estimation and validation. *IEEE Trans Biomed Eng*, **42**(7), 658–665.

Pasley BN and Freeman RD. (2008). Neurovascular coupling. *Scholarpedia*, **3**(3), 5340, revision 91581. revision 91581.

Persson J, Kalpouzos G, Nilsson LG, Ryberg M and Nyberg L. (2011). Preserved hippocampus activation in normal aging as revealed by fmri. *Hippocampus*, **21**(7), 753–766.

Petridou N, Caballero Gaudes C, Dryden IL, Francis ST and Gowland PA. (2013). Periods of rest in fmri contain individual spontaneous events which are related to slowly fluctuating spontaneous activity. *Human Brain Mapp*, **34**(6), 1319–1329.

Polman CH, Reingold SC, Edan G, Filippi M, Hartung HP, Kappos L, Lublin FD, Metz LM, McFarland HF, O'Connor PW, Sandberg-Wollheim M, Thompson AJ, Weinshenker BG and Wolinsky JS. (2005). Diagnostic criteria for multiple sclerosis: 2005 revisions to the "McDonald Criteria". *Ann Neurol*, **58**(6), 840–846.

Portilla J and Simoncelli EP. (2000). A parametric texture model based on joint statistics of complex wavelet coefficients. *Int J Comput Vision*, **40**, 49–70.

Power JD, Barnes KA, Snyder AZ, Schlaggar BL and Petersen SE. (2012). Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. *Neuroimage*, **59**(3), 2142–2154.

Preti MG, Leonardi N, Karahanoglu FI, Grouiller F, Genetti M, Seeck M, Vulliemoz S and Van De Ville D. (2014). Epileptic network activity revealed by dynamic functional connectivity in simultaneous EEG-fMRI. In *Proc. IEEE International Symposium on Biomedical Imaging (ISBI)*, 9–12, Beijing, China.

Raichle ME, MacLeod AM, Snyder AZ, Powers WJ, Gusnard DA and Shulman GL. (2001). A default mode of brain function. *Proc Natl Acad Sci U S A*, **98**(2), 676–682.

Raj A, Kuceyeski A and Weiner M. (2012). A network diffusion model of disease progression in dementia. *Neuron*, **73**(6), 1204–1215.

Ram I, Elad M and Cohen I. (2011). Generalized tree-based wavelet transform. *IEEE Trans Signal Process*, **59**, 4199 –4209.

Richiardi J, Achard S, Bunke H and Van de Ville D. (2013). Machine learning with brain graphs: Predictive modeling approaches for functional imaging in systems neuroscience. *IEEE Signal Processing Mag*, **30**(3), 58–70.

Richiardi J, Eryilmaz H, Schwartz S, Vuilleumier P and Van De Ville D. (2011). Decoding brain states from fMRI connectivity graphs. *NeuroImage*, **56**(2), 616–626.

Richiardi J, Van De Ville D and Eryilmaz H. (2010). Low-dimensional embedding of functional connectivity graphs for brain state decoding. In *Proc IEEE International Workshop on Pattern Recognition in Neuroimaging (PRNI)*, 21 –24.

Richiardi J, Gschwind M, Simioni S, Annoni JM, Greco B, Hagmann P, Schluep M, Vuilleumier P and Van De Ville D. (2012). Classifying minimally disabled multiple sclerosis patients from resting state functional connectivity. *NeuroImage*, **62**(3), 2021–2033.

Rocca MA, Valsasina P, Absinta M, Riccitelli G, Rodegher ME, Misci P, Rossi P, Falini A, Comi G and Filippi M. (2010). Default-mode network dysfunction and cognitive impairment in progressive MS. *Neurology*, **74**(16), 1252–1259.

Rocca MA, Colombo B, Falini A, Ghezzi A, Martinelli V, Scotti G, Comi G and Filippi M. (2005). Cortical adaptation in patients with MS: a cross-sectional functional MRI study of disease phenotypes. *Lancet Neurol*, **4**(10), 618–626.

Rocca MA, Valsasina P, Martinelli V, Misci P, Falini A, Comi G and Filippi M. (2012). Large-scale neuronal network dysfunction in relapsing-remitting multiple sclerosis. *Neurology*, **79**(14), 1449–1457.

Romero J, Alexander S, Baid S, Jain S and Papadakis M. (2009). The geometry and the analytic properties of isotropic multiresolution analysis. *Adv Comp Math*, **31**, 283–328.

Roosendaal SD, Schoonheim MM, Hulst HE, Sanz-Arigita EJ, Smith SM, Geurts JJG and Barkhof F. (2010). Resting state networks change in clinically isolated syndrome. *Brain*, **133**(Pt 6), 1612–1621.

Rugg MD and Vilberg KL. (2013). Brain networks underlying episodic memory retrieval. *Curr Opin Neurobiol*, **23**(2), 255–260.

Sadaghiani S, Hesselmann G and Kleinschmidt A. (2009). Distributed and antagonistic contributions of ongoing activity fluctuations to auditory stimulus detection. *J Neurosci*, **29**(42), 13410–13417.

Sadaghiani S, Scheeringa R, Lehongre K, Morillon B, Giraud AL and Kleinschmidt A. (2010). Intrinsic connectivity networks, alpha oscillations, and tonic alertness: a simultaneous electroencephalography/functional magnetic resonance imaging study. *J Neurosci*, **30**(30), 10243–10250.

Salvador R, Suckling J, Schwarzbauer C and Bullmore E. (2005). Undirected graphs of frequency-dependent functional connectivity in whole brain networks. *Philos Trans R Soc Lond B Biol Sci*, **360** (1457), 937–946.

Sandryhaila A and Moura JMF. (2013). Discrete signal processing on graphs. *IEEE Trans Signal Process*, **61**, 1644–1656.

Scarth G, McIntyre M, Wowk B and Somorjai R. (1995). Detection of novelty in functional images using fuzzy clustering. In *Society of Magnetic Resonance (SMR)*, 238, Nice.

Schaefer A, Margulies DS, Lohmann G, Gorgolewski KJ, Smallwood J, Kiebel SJ and Villringer A. (2014). Dynamic network participation of functional connectivity hubs assessed by resting-state fmri. *Front Hum Neurosci*, **8**, 195.

Schölkopf B, Smola A and Müller KR. (1998). Nonlinear component analysis as a kernel eigenvalue problem. *Neural Comput*, **10**(5), 1299–1319.

Schoonheim MM and Filippi M. (2012). Functional plasticity in MS: friend or foe? *Neurology*, **79**(14), 1418–1419.

Seeley WW, Crawford RK, Zhou J, Miller BL and Greicius MD. (2009). Neurodegenerative diseases target large-scale human brain networks. *Neuron*, **62**(1), 42–52.

Seeley WW, Menon V, Schatzberg AF, Keller J, Glover GH, Kenna H, Reiss AL and Greicius MD. (2007). Dissociable intrinsic connectivity networks for salience processing and executive control. *J Neurosci*, **27**(9), 2349–2356.

Shen HW, Cheng XQ, Wang YZ and Chen Y. (2012). A dimensionality reduction framework for detection of multiscale structure in heterogeneous networks. *Journal of Computer Science and Technology*, **27** (2), 341–357.

Shirer WR, Ryali S, Rykhlevskaia E, Menon V and Greicius MD. (2012). Decoding subject-driven cognitive states with whole-brain connectivity patterns. *Cereb Cortex*, **22**(1), 158–165.

Shulman GL, Fiez JA, Corbetta M, Buckner RL, Miezin FM, Raichle ME and Petersen SE. (1997). Common blood flow changes across visual tasks: Ii. decreases in cerebral cortex. *J Cogn Neurosci*, **9** (5), 648–663.

Shuman D, Narang S, Frossard P, Ortega A and Vandergheynst P. (2013). The emerging field of signal processing on graphs: Extending high-dimensional data analysis to networks and other irregular domains. *IEEE Signal Process Mag*, **30**(3), 83–98.

Smith SM, Fox PT, Miller KL, Glahn DC, Fox PM, Mackay CE, Filippini N, Watkins KE, Toro R, Laird AR and Beckmann CF. (2009). Correspondence of the brain's functional architecture during activation and rest. *Proc Natl Acad Sci U S A*, **106**(31), 13040–13045.

Smith SM, Miller KL, Moeller S, Xu J, Auerbach EJ, Woolrich MW, Beckmann CF, Jenkinson M, Andersson J, Glasser MF, Van Essen DC, Feinberg DA, Yacoub ES and Ugurbil K. (2012). Temporally-independent functional modes of spontaneous brain activity. *Proc Natl Acad Sci U S A*, **109**(8), 3131–3136.

Sowell ER, Peterson BS, Thompson PM, Welcome SE, Henkenius AL and Toga AW. (2003). Mapping cortical change across the human life span. *Nat Neurosci*, **6**(3), 309–315.

Spencer M, Daly I and Nasuto S. (2013). Meso-scale models of evolving networks for EEG analysis. In *Workshop on Time-Dependent and Multiplex Networks*, Oxford, UK.

Sporns O. (2010). *Networks of the brain*. MIT Press, Cambridge, MA.

Sporns O. (2013). Network attributes for segregation and integration in the human brain. *Curr Opin Neurobiol*, **23**(2), 162–171.

Spreng RN, Stevens WD, Chamberlain JP, Gilmore AW and Schacter DL. (2010). Default network activity, coupled with the frontoparietal control network, supports goal-directed cognition. *NeuroImage*, **53** (1), 303–317.

Starck T, Nikkinen J, Remes J, Rahko J, Moilanen I, Tervonen O and Kiviniemi V. (2012). Temporally varying connectivity between ICA default-mode sub-networks - ASD vs. controls. In *Proc. Organization for Human Brain Mapping (OHBM)*, 1, Beijing, China.

Strang G. (1999). The discrete cosine transform. *SIAM Rev.*, **41**(1), 135–147.

Szell M, Lambiotte R and Thurner S. (2010). Multirelational organization of large-scale social networks in an online world. *Proc Natl Acad Sci. USA*, **107**(31), 13636–13641.

Sämann PG, Wehrle R, Hoehn D, Spoormaker VI, Peters H, Tully C, Holsboer F and Czisch M. (2011). Development of the brain's default mode network from wakefulness to slow wave sleep. *Cereb Cortex*, **21**(9), 2082–2093.

Tagliazucchi E, Balenzuela P, Fraiman D and Chialvo DR. (2012). Criticality in large-scale brain fmri dynamics unveiled by a novel point process analysis. *Front Physiol*, **3**, 15.

Theiler J, Eubank S, Longtin A, Galdrikian B and Farmer JD. (1992). Testing for nonlinearity in time series: the method of surrogate data. *Physica D: Nonlinear Phenomena*, **58**, 77 – 94.

Thompson GJ, Magnuson ME, Merritt MD, Schwarb H, Pan WJ, McKinley A, Tripp LD, Schumacher EH and Keilholz SD. (2013a). Short-time windows of correlation between large-scale functional brain networks predict vigilance intraindividually and interindividually. *Hum Brain Mapp*, **34**(12), 3280–3298.

Thompson GJ, Magnuson ME, Merritt MD, Schwarb H, Pan WJ, McKinley A, Tripp LD, Schumacher EH and Keilholz SD. (2013b). Short-time windows of correlation between large-scale functional brain networks predict vigilance intraindividually and interindividually. *Human Brain Mapp*, **34** (12), 3280–3298.

Torrence C and Webster PJ. (1999). Interdecadal changes in the ENSO-monsoon system. *J. Climate*, **12** (8), 2679–2690.

Tremblay N and Borgnat P. (2013). Multiscale community mining in networks using the graph wavelet transform of random vectors. In *Proc IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, 463–466, Atlanta, GA.

Tripathi A, Klami A and Kaski S. (2008). Simple integrative preprocessing preserves what is shared in data sources. *BMC Bioinformatics*, **9**(1), 1–13.

Tucker L. (1966). Some mathematical notes of three-mode factor analysis. *Psychometrika*, **31**, 279–311.

Tzourio-Mazoyer N, Landeau B, Papathanassiou D, Crivello F, Etard O, Delcroix N, Mazoyer B and Joliot M. (2002). Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *NeuroImage*, **15**(1), 273–289.

Van Cauter E, Leproult R and Plat L. (2000). Age-related changes in slow wave sleep and rem sleep and relationship with growth hormone and cortisol levels in healthy men. *JAMA*, **284**(7), 861–868.

Van De Ville D, Blu T and Unser M. (2004). Integrated wavelet processing and spatial statistical testing of fMRI data. *NeuroImage*, **23**, 1472–1485.

Van de Ville D, Britz J and Michel CM. (2010). EEG microstate sequences in healthy humans at rest reveal scale-free dynamics. *Proc Natl Acad Sci U S A*, **107**(42), 18179–18184.

Van De Ville D, Seghier ML, Lazeyras F, Blu T and Unser M. (2007). WSPM: wavelet-based statistical parametric mapping. *Neuroimage*, **37**(4), 1205–1217.

Varoquaux G, Sadaghiani S, Pinel P, Kleinschmidt A, Poline JB and Thirion B. (2010). A group model for stable multi-subject ICA on fMRI datasets. *Neuroimage*, **51**(1), 288–299.

Vasilescu MAO and Terzopoulos D. (2002). Multilinear analysis of image ensembles: Tensorfaces. In *Proc European Conference on Computer Vision (ECCV)*, 447–460, Copenhagen.

Vasilescu M. (2011). Multilinear projection for face recognition via canonical decomposition. In *Proc IEEE Automatic Face Gesture Recognition and Workshops*, 476–483, Santa Barbara, CA.

Vincent JL, Patel GH, Fox MD, Snyder AZ, Baker JT, Van Essen DC, Zempel JM, Snyder LH, Corbetta M and Raichle ME. (2007). Intrinsic functional architecture in the anaesthetized monkey brain. *Nature*, **447**(7140), 83–86.

Vincent JL, Snyder AZ, Fox MD, Shannon BJ, Andrews JR, Raichle ME and Buckner RL. (2006). Coherent spontaneous activity identifies a hippocampal-parietal memory network. *J Neurophysiol*, **96**(6), 3517–3531.

Von Below J. (1985). A characteristic equation associated to an eigenvalue problem on c^2-networks. *Linear Algebra Appl*, **71**, 309—-325.

Waites AB, Stanislavsky A, Abbott DF and Jackson GD. (2005). Effect of prior cognitive state on resting state networks measured with functional connectivity. *Hum Brain Mapp*, **24**(1), 59–68.

Walsh V and Cowey A. (2000). Transcranial magnetic stimulation and cognitive neuroscience. *Nat Rev Neurosci*, **1**(1), 73–79.

Wang L, Laviolette P, O'Keefe K, Putcha D, Bakkour A, Van Dijk KRA, Pihlajamäki M, Dickerson BC and Sperling RA. (2010). Intrinsic connectivity between the hippocampus and posteromedial cortex predicts memory performance in cognitively intact older individuals. *NeuroImage*, **51**(2), 910–917.

Worsley KJ, Marrett S, Neelin P and Friston KJ. (1996). A unified statistical approach for determining significant signals in images of cerebral activation. *Hum Brain Mapp*, **4**, 58–73.

Wu GR, Liao W, Stramaglia S, Ding JR, Chen H and Marinazzo D. (2013). A blind deconvolution approach to recover effective connectivity brain networks from resting state fmri data. *Med Image Anal*, **17**(3), 365–374.

Yourganov G, Chen X, Lukic AS, Grady CL, Small SL, Wernick MN and Strother SC. (2011). Dimensionality estimation for optimal detection of functional networks in BOLD fMRI data. *NeuroImage*, **56** (2), 531–543.

Zhang D and Raichle ME. (2010). Disease and the brain's dark energy. *Nat Rev Neurol*, **6**(1), 15–28.

Zlokovic BV and Apuzzo ML. (1998). Strategies to circumvent vascular barriers of the central nervous system. *Neurosurgery*, **43**(4), 877–878.