DTU Library

Brain Network Modelling

Andersen, Kasper Winther

Publication date:
2014

Document Version
Publisher's PDF, also known as Version of record

[Link back to DTU Orbit](#)

Citation (APA):
Andersen, K. W. (2014). *Brain Network Modelling*. Technical University of Denmark. DTU Compute PHD-2014 No. 326

Brain Network Modelling

Kasper Winther Andersen

To Laura & Mads-Emil

Summary

Three main topics are presented in this thesis. The first and largest topic concerns network modelling of functional Magnetic Resonance Imaging (fMRI) and Diffusion Weighted Imaging (DWI). In particular nonparametric Bayesian methods are used to model brain networks derived from resting state fMRI data. The models used are the Infinite Relational Model (IRM), Bayesian Community Detection (BCD), and Infinite Diagonal Model (IDM). The models have different constraints on how they cluster nodes. IRM is flexible in the sense that it allows for complex interactions between clusters of nodes. BCD conforms to the definition of *community structure* in the sense that it forces clusters of nodes to have larger density of internal connections than external connections. IDM models only the linking within a cluster and treats linking between clusters as background noise. The models are evaluated for their ability to reproduce node clustering and predict unseen data. Comparing the models on whole brain networks, BCD and IRM showed better reproducibility and predictability than IDM, suggesting that resting state networks exhibit community structure. This also points to the importance of using models, which allow for complex interactions between all pairs of clusters. In addition, it is demonstrated how the IRM can be used for segmenting brain structures into functionally coherent clusters.

A new nonparametric Bayesian network model is presented. The model builds upon the IRM and can be used to infer shared clustering structure across different types of networks. The model is used to jointly model fMRI and DWI networks. However, results show only a limited amount of sharing across fMRI and DWI networks. Using the model within the same modality can reveal the clustering consistency across scans. A high consistency was found between DWI networks and an intermediate level of consistency was found between fMRI net-

works. The model is of interest for other applications, for instance in finding *dissimilarity* between network structure in case-control studies.

The second topic of the thesis concerns local functional connectivity. In particular the local functional connectivity is studied in patients with multiple sclerosis (MS). The functional connectivity in a small neighborhood was estimated using Kendall's Coefficient of Concordance (KCC). By generating voxelwise KCC maps, MS patients were compared with healthy controls. MS patients had reduced KCC in cerebellum and KCC correlated negatively with disease progression. Lesion load of the left cerebellar peduncles correlated negatively with KCC suggesting that the reduced local connectivity in MS is caused by disrupted inputs to the cerebellum.

The final topic of this thesis concerns model selection for Gaussian Kernel Principal Component Analysis (KPCA) denoising. KPCA can be used for non-linear denoising by mapping data to feature space using a non-linear map. By projecting data onto a subspace in feature space and mapping this projection back to input space noise in data is (hopefully) removed. However, two important parameters must be set, namely the scale of the Gaussian kernel and the subspace dimensionality. A principled method for selecting these two parameters is presented. The method is based on maximizing the signal energy in feature space. When testing on synthetic and real data, the method outperformed a number of other heuristics in terms of signal to noise ratio of the denoised data.

Resumé (Danish)

Den menneskelige hjerne består af et gigantisk netværk af forbundne neuroner, der i samarbejde former vores eksistens og lader os agere ud fra sanseindtryk. Hjerneaktivitet kan måles med funktionel magnetisk resonans skanninger (fMRI) og ved at bruge matematiske netværksmodeller kan hjerneområder grupperes ud fra hvordan de arbejder sammen. I denne afhandling bliver forskellige modeller, der kan danne disse grupperinger, testet for deres stabilitet og prædiktionsevne. Resultatet viser at en model, der favoriserer *community* struktur, klarer sig bedst. *Community* struktur er karakteriseret ved høj link-tæthed indenfor en gruppe og lav link-tætheder til andre grupper og findes ofte i fx sociale netværk. Ydermere bliver en model, der kan finde fælles grupperinger i forskellige netværks typer, udviklet. Modellen anvendes på funktionelle og strukturelle hjernenetværk. Resultatet viser dog en beskeden fælles struktur mellem de to typer netværk, hvilket dog reflekterer begrænsninger i optagelsen af data. Brugen af disse netværksmodeller til analyse af hjernenetværk er relativ ny, men modellerne har et stort potentiale til bl.a. at finde forskelle i hjernens kommunikation mellem raske og syge.

I et andet studie inkluderet i afhandlingen undersøges lokale netværksegenskaber i funktionelle netværk. Den lokale konnektivitet, som er et mål for hvor godt et lille område af hjernen er funktionelt forbundet med de omkringliggende hjerneområder, bliver sammenlignet mellem patienter med multipel sklerose (MS) og raske personer. Resultatet viser en nedsat lokal konnektivitet i lillehjernen hos MS, og ydermere at konnektiviteten bliver lavere ved højere sygdomsgrad. Det tyder på at dette skyldes læsioner i hjerneområdet, der rummer forbindelserne til lillehjernen.

Det sidste område som berøres i afhandlingen er vedrørende fjernelse af støj fra data. Kernel Principal Component Analysis (KPCA) kan bruges til støjfjernelse men resultatet afhænger af valg af parametre. Vi har udviklet en metode hvorved udvælgelsen af disse parametre sker automatisk og vi viser at metoden klarer sig bedre end andre heuristikker til parametervalg.

Preface

This thesis was prepared partly at Department of Applied Mathematics and Computer Science, Technical University of Denmark and partly at Danish Research Centre for Magnetic Resonance, Centre for Functional and Diagnostic Imaging and Research, Copenhagen University Hospital (DRCMR) in partial fulfillment of the requirements for acquiring the Ph.D. degree in engineering. The project was funded partly ($\frac{2}{3}$) by a research grant from the Lundbeck Foundation given to Hartwig R. Siebner and partly ($\frac{1}{3}$) by DTU Compute. My main supervisor was Professor Lars Kai Hansen, Department of Applied Mathematics and Computer Science, Technical University of Denmark. In addition I had three co-supervisors; Professor Hartwig R. Siebner, Ph.D. Kristoffer Hougaard Madsen, and Ph.D. Tim B. Dyrby, all from DRCMR. The thesis consists of a summary report and a collection of two journal papers, a book chapter, a conference paper, a journal paper under review, and a paper in preparation. The work was carried out between December 2010 and January 2014.

Lyngby, January 28, 2014

Kasper Winther Andersen

Publication List

Papers included in this thesis

Journal Papers

[A] Jørgensen, K. W., and Hansen, L. K. (2012), 'Model Selection for Gaussian Kernel PCA Denoising', *Neural Networks and Learning Systems, IEEE Transactions on*, 23(1), 163-168.
Note that this paper was published under my former lastname Jørgensen.

[B] Dogonowski, A.-M.*, Andersen, K. W.*, Madsen, K. H., Sørensen, P. S., Paulson, O. B., Blinkenberg, M., and Siebner, H. R. (2014), 'Multiple sclerosis impairs regional functional connectivity in the cerebellum', *NeuroImage: Clinical*, 4, 130-138.
* These authors contributed equally to the manuscript.

Conference Proceedings

[C] Andersen, K. W., Mørup, M., Siebner, H. R., Madsen, K. H., and Hansen, L. K. (2012), 'Identifying modular relations in complex brain networks', in '2012 IEEE International Workshop on Machine Learning for Signal Processing', IEEE.

Book Chapter

[D] Andersen, K. W., Madsen, K. H., Siebner, H. R., Hansen, L. K., and Mørup, M. (2012), 'Identification of Functional Clusters in the Striatum Using Infinite Relational Modeling', in G. Langs, I. Rish, M. Grosse-Wentrup and B. Murphy, eds, 'Machine Learning and Interpretation in

Neuroimaging', Lecture Notes in Computer Science, Springer Berlin Heidelberg, pp. 226-233.

Submitted

[E] Andersen, K. W., Madsen, K. H., Siebner, H. R., Schmidt, M. N., Mørup, M., and Hansen, L. K. (2013), 'Coordinated activation in the resting brain is community structured', *under review.*

In preparation

[F] Andersen, K. W., Herlau, T., Mørup, M., Schmidt, M. N., Madsen, K. H., Dyrby, T. D., Lyksborg, M., Siebner, H. R., and Hansen, L. K., (2014) 'Joint Modelling of Structural and Functional Brain Networks', *in preparation.*

Material not included in this thesis

Journal Paper

- Fryer, S. L., Jorgensen, K. W., Yetter, E. J., Daurignac, E. C., Watson, T. D., Shanbhag, H., Krystal J. H., Mathalon, D. H. (2012), 'Differential brain response to alcohol cue distractors across stages of alcohol dependence', Biological psychology, 92(2), 282-291.

Workshop Paper

- Andersen, K. W., Herlau, T., Mørup, M., Schmidt, M. N., Madsen, K. H., Dyrby, T. D., Lyksborg, M., Siebner, H. R., Hansen, L. K., (2012) 'Joint Modelling of Structural and Functional Brain Networks', Machine Learning and Interpretation in Neuroimaging workshop, Lake Tahoe, USA, 2012.

Conference Abstract

- Dogonowski, A.-M., Andersen K. W., Madsen K. H., Sørensen, P. S., Paulson, O. B., Blinkenberg, M., and Siebner, H. R., 'Multiple sclerosis impairs regional functional connectivity in the cerebellum', *28th Congress of the European Committee for Treatment and Research in Multiple Sclerosis*, Lyon, France, 2012.

In preparation

- Nilsson, S. J., Andersen, K. W., Raffin, E., van Nuenen, B. F. L., Kuhtz-Bushbeck, J., and Siebner, H. R. (2014), 'Dorsal premotor cortex codes for optimal performance during a grasp-and-lift task'. *In preparation*

Software

- KPCA Matlab toolbox
  `http://www2.imm.dtu.dk/pubdb/views/publication_details.php?id=6236`

- fMRI IRM Matlab toolbox
  `https://brainconnectivity.compute.dtu.dk/?download=85`

Acknowledgements

Nomenclature

Abbreviations

AAL	Automated Anatomical Labelling - brain atlas
AUC	Area Under Curve
BOLD	Blood-Oxygen Level Dependent
BCD	Bayesian Community Detection
CRP	Chinese Restaurant Process
CST	Cortico-Spinal Tract
DRCMR	Danish Research Centre for Magnetic Resonance
DWI	Diffusion Weighed Imaging
fMRI	Functional Magnetic Resonance Imaging
FSL	FMRIB Software Library
HC	Healthy Controls
HRF	Hemodynamic Response Function
IDM	Infinite Diagonal Model
IHW	Infinite Hofman-Wiggins
IRM	Infinite Relational Model
KCC	Kendall's Coefficient of Concordance
kPA	Kernel Parallel Analysis
KPCA	Kernel Principal Component Analysis
LL	Lesion Load
MCMC	Markov Chain Monte Carlo
MI	Mutual Information
NMI	Normalized Mutual Information
MS	Multiple Sclerosis

NPAIRS	Nonparametric Prediction, Activation, Influence, and Reproducibility resampling
PA	Parallel Analysis
PCA	Principal Component Analysis
ROC	Receiver Operating Characteristic
rs-fMRI	Resting-State Functional Magnetic Resonance Imaging
SPM	Statistical Parametric Mapping

Notation

y	scalar
$\mathbf{y}$	vector
y_i	i-th element of $\mathbf{y}$
$\mathbf{Y}$	matrix
Y_{ij}	(i, j)-th element of $\mathbf{Y}$
$(\cdot)^\top$	transpose
$\bar{\mathbf{y}}$	mean

Fixed symbols

Nonparametric Bayesian network analysis

$\mathbf{A}$	adjacency matrix
$\mathbf{A}^{(n)}$	adjacency matrix for subject n
$\tilde{\mathbf{A}}$	the set of adjacency matrices $\{\mathbf{A}^{(i)}\}_{i=1}^{N}$
J	number of nodes
K	number of clusters
N_{kl}^{+}	number of links across all subjects between cluster k and l
N_{kl}^{-}	number of nonlinks across all subjects between cluster k and l
N_{b}^{+}	number of links across all subjects which fall between clusters
N_{b}^{-}	number of nonlinks across all subjects which fall between clusters
N_{w}^{+}	number of links across all subjects which fall within clusters
N_{w}^{-}	number of nonlinks across all subjects which fall within clusters

n_k	number of nodes assigned to cluster k
r_{il}^{+}	number of links from node i to all nodes in cluster l
r_{il}^{-}	number of nonlinks from node i to all nodes in cluster l
v	hyper-parameter for cluster-gap prior
$\mathbf{z}$	cluster assignment vector
$\mathbf{z}_{\setminus i}$	assignments of all nodes except node i
α	hyper-parameter for the CRP
β	hyper-parameter for link probability prior
$\boldsymbol{\rho}$	cluster link probability matrix
$\dot{\boldsymbol{\rho}}$	set of within-cluster link probabilities
$\ddot{\boldsymbol{\rho}}$	set of between-cluster link probabilities
γ	cluster gap used in BCD

KPCA

$\mathbf{a}_i$	eigenvector of kernel matrix
$\mathbf{C}$	covariance matrix in feature space
$\mathcal{F}$	feature space
$\mathbf{G}$	kernel matrix
$g(\cdot,\cdot)$	kernel function
$\mathbf{u}_i$	i-th eigenvector in feature space
$\mathcal{X}$	input space
λ_i	i-th eigenvalue
φ	map from input space to feature space
σ	scale of the Gaussian kernel

Contents

CHAPTER 1

Introduction

From the brain and the brain alone arise our pleasures, joys, laughter and jests, as well as our sorrows, pains and griefs.
- Hippocrates

Each of the 10^{11} neurons in the human brain has on average 7000 synaptic connections constituting a gigantic neural network. The communication between different brain regions is forming our memories, experiences and guides us to perform actions. The neural network constitutes one of the most intriguing networks, and a huge amount of effort is put into understanding its details – from a behavioral level to the micro-structural level. By understanding the brain's function and structure we can better understand brain diseases and help remedy pathologies. Enormous effort is put into understanding how brain regions are connected, e.g., the newly founded *Human Connectome Project* aims at mapping the so-called connectome, an exhaustive map of the neural connections.

Resting state functional magnetic resonance imaging (rs-fMRI) has become one of the most widely used methods to study brain function. The first study which attributed the resting fMRI signal with neural activity was made by Biswal et al. (1995) when they correlated the time series of voxels in one part of the motor cortex and observed that other regions of the motor network showed significant correlations. Since then, the number of studies using resting state as a research tool has rapidly increased. In fact, searching PubMed (`www.pubmed.org`) for resting state fMRI entries reveals an exponential increasing number of entries over the last decade, whereas fMRI entries in general only increase linearly (Figure 1.1).

The classical way of analyzing task fMRI is to correlate the fMRI signal with regressors describing the task, however, in rs-fMRI there are no external stimuli or task. Therefore the common strategy is to use correlation type analysis, which aims at exploring the mutual signal fluctuations between different brain regions and thereby infer functional connectivity. This is commonly done by correlating the time series from a seed region with the rest of the brain to explore the connectivity pattern or to use multivariate methods which explores common spatio-temporal features of the BOLD signal. Independent Component Analysis is one of the most popular of such methods and has revealed that a number of so-called resting state networks exist (Damoiseaux et al., 2006). These networks have shown to be very consistent across subjects and represent brain regions which fluctuate in synchrony at a relatively low frequency (in the order of 0.1-0.01Hz). Alterations in the representations of resting state networks have been found in a wide range of neurological and psychiatric diseases such as schizophrenia (Skudlarski et al., 2010), Alzheimer's disease (Sorg et al., 2007), and Parkinson's disease (Wu, Wang, Chen, Zhao, Li and Chan, 2009).

Diffusion Weighted Imaging (DWI) is a popular tool for estimating structural connectivity. DWI measures the movement of water molecules and can thereby reveal the local directions of fiber bundles. By mapping fiber directions in a fine grid it is possible to use tractography methods to follow the directions from a seed and thereby estimate how brain regions are anatomically connected.
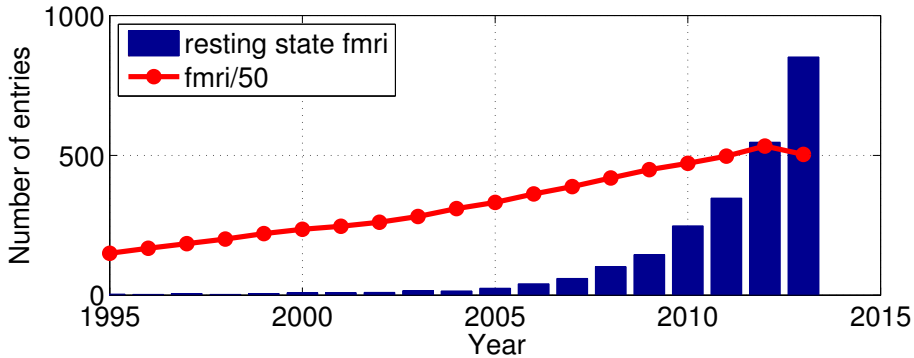
Figure 1.1: Number of entries per year in PubMed (`www.pubmed.org`) searches using the search queries 'fmri' and '("resting state") AND fmri' respectively. The number of `resting state fmri` entries are increasing exponentially while the `fmri` entries only increase linearly. Note that the number of fmri entries has been divided by 50. The search was performed on January 23, so the numbers for 2013 might not be fully updated. Figure adapted from (Snyder and Raichle, 2012).

A natural framework for analyzing the structural and functional connections is to use graph theory. In graph theory, connections between nodes are characterized, and different network properties can therefore be investigated. One of such properties, which is of particular interest for this thesis, is how nodes group together to form so-called *communities* (Newman and Girvan, 2004). A community is a group of nodes with a high link density within the community with low link density to the rest of the network. Forming such communities in brain networks can reveal how different parts of the brain work together. The main part of the thesis is using a family of nonparametric Bayesian methods for finding node clustering in networks. Nonparametric Bayesian models are models which cannot be parameterized by a finite number of parameters, as the model is allowed to grow with the complexity of the data. The models put different constrains on how clusters relate to each other. The most flexible model, the Infinite Relational Model (IRM)(Kemp et al., 2006; Xu et al., 2006), have no constrains on how clusters can relate to each other and therefore allow for a flexible node clustering. The Bayesian Community Detection (BCD) (Mørup and Schmidt, 2012) adheres to the definition of community structure as it forces a stronger linking within a community than to other communities. However, BCD still allows for complex relations to other clusters. These models are compared with two other alternatives which treat all linking between clusters as background noise.

Another topic of the thesis is modelling of connectivity at the local level. Here,

Kendall's coefficient of concordance is used to estimate the synchrony of fMRI signals in small neighborhoods reflecting the local connectivity properties. Maps representing local connectivity are generated and compared between a group of multiple sclerosis (MS) patients and healthy controls (HC).

The final topic of the thesis is concerning model selection for Gaussian Kernel Principal Component Analysis (KPCA) denoising. In Gaussian KPCA the two parameters must be set, namely the Gaussian scale σ and the subspace dimension in feature space. A principled method for simultaneously selecting these parameters is presented.

1.1   Organization of this Thesis

The thesis is organized as follows.

Chapter 2 describes the fMRI and DWI datasets used.

Chapter 3 describes the efforts done on utilizing Bayesian complex network models for estimating brain connectivity at the whole brain level. The methods are introduced and the paper contributions are summarized.

Chapter 4 describes work on estimating functional connectivity at the local level and how the local connectivity is altered in patients with MS.

Chapter 5 introduces KPCA and shows how it can be used for data denoising. In particular the contribution on selecting parameters for Gaussian KPCA is described.

Chapter 6 summarizes and concludes on the thesis as a whole.

Data

Not everything that can be counted counts, and not everything that counts can be counted.
- Albert Einstein

This chapter will first introduce functional Magnetic Resonance Imaging (fMRI). Then the datasets, which are used in this thesis, will be described along with the preprocessing steps applied to the data.

2.1 Functional Magnetic Resonance Imaging

In fMRI the blood-oxygen level dependent (BOLD) signal is measured (Ogawa et al., 1990). The BOLD signal is an in-direct measure of the underlying neuronal activity. When neurons are activated they consume oxygen in the nearby blood, which then becomes deoxygenated. Deoxyhemoglobin is paramagnetic and therefore distorts the local magnetic field, which is detectable by a MR scanner. The increased energy demand causes an increase in the local cerebral blood flow, resulting in an increased oxygen-to-deoxygen ratio, which is observed via an increased MR signal response. This increased MR response due to increased neural activity is commonly known as the hemodynamic response function (HRF). The HRF is a relatively slow response and peaks at around 5 seconds after stimulation and returns to baseline again at 12-20 seconds post stimuli.

2.2 DRCMR Data

The data described in this section was collected at the Danish Research Centre for Magnetic Resonance (DRCMR) by Anne-Marie Dogonowski and Kristoffer H. Madsen.

2.2.1 Subjects

The dataset from DRCMR consist of 42 MS patients and 30 HC. All MS patients fulfilled the revised McDonald criteria (Polman et al., 2005) and the HC had no history of neurological or psychiatric diseases. Subjects were tested with the Edinburgh handedness inventory (Oldfield, 1971), which revealed that 27 HC and 39 MS subjects were right-handed. All the patients were recruited from the Danish Multiple Sclerosis Center, which is a large MS center in Copenhagen, Denmark. Of the patients 27 were relapsing-remitting MS (RR-MS) and 15 patients were secondary progressive (SP-MS). The patients did not have any relapse 3 months prior to scanning, and they were all rated with the Expanded Disability Status Scale (EDSS) (Kurtzke, 1983). The patients' EDSS scores ranged from 0 to 7 (median 4.3). Please see Anne-Marie Dogonowski's PhD Thesis for further details on the subjects (Dogonowski, 2012).

2.2.2 Resting State Data

The MRI was conducted on a 3T Siemens (Erlangen, Germany) Magnetom Trio scanner. The first scan during in the session was a 20 minutes resting-state fMRI (T2* weighted echo planar imaging) scan. The subjects were instructed to rest (eyes closed) and not think of anything in particular. Subjects were also asked to avoid any voluntary movements. The scan parameters were: TR=2490ms, TE=30ms, flip-angle=90°, 480 volumes, 42 axial slices, 3 mm isotropic resolution, field of view (FOV) 192x192mm. Cardiac cycles were recorded using infrared pulse oximeter and respiratory cycles were recorded using pneumatic thoracic belt. After the functional scan a magnetization prepared rapid gradient echo (MPRAGE) scan was conducted using the scan parameters; TR=1550ms, TE=3.04ms, inversion time (IT) = 800ms, 192 sagittal slices, 1 mm isotropic resolution, FOV=256mm, flip-angle = 9°.

Resting State Data Pre-processing
Before further analysis the fMRI data was preprocessed using statistical parametric mapping (SPM8, Wellcome Trust Centre for Neuroimaging[1]) implemented in Matlab 7.9 (MathWorks, Massachusetts, USA). The first five volumes were discarded to reach steady-state magnetization. The remaining images were realigned using a two-step procedure, which first aligns all images to the first volume in the time series and then to the mean volume. The EPI images were then normalized to standard space by first co-registering them to the T1-image using a 6-parameter rigid-body transformation. The T1 image were normalized to MNI305 space by using the unified segmentation/normalisation procedure (Ashburner and Friston, 2005), which is implemented in SPM. The EPI images was then normalised by applying the same normalisation transformation.

The EPI images are confounded by noise induced both by scanner instabilities and physiological noise (cardiac, pulsation and respiration) (Dagli et al., 1999; Glover et al., 2000; Lund et al., 2006; Smith et al., 1999). These different noise contribution can give rise to signal changes similar to those observed in rs-fMRI (Birn et al., 2006) and therefore the time series were temporally filtered prior to further analysis. A linear filter were constructed which included the following components. A 1/128 Hz cut-off high-pass filter, implemented as discrete cosine basis functions, was included to remove low-frequency scanner drifts. The cardiac and respiratory signals were modelled using a Fourier expansion of the cardiac (10 parameters) and respiratory (6 parameters) cycles and the first order cardiac by respiration interaction effects (4 parameters) (Glover et al., 2000). Head motion was modelled by a Taylor expansion of the realignment parameters and their time-lagged version (Friston et al., 1996). Birn et al. (2006) showed that the respiration volume over time produce signal changes similar to rs-fMRI

---

[1]http://www.fil.ion.ucl.ac.uk/spm

signals. To account for this the filter included time-lagged versions of the respiration volume (from -20 to 20 seconds in 1 second intervals). Finally, the filter included time series from voxels in the cerebrospinal fluid and white matter.

2.2.3 Diffusion Weighted Imaging

Each subject underwent two diffusion weighted imaging (DWI) scans using a twice-refocused spin echo sequence (Reese et al., 2003). For each session, diffusion along 61 directions were recorded with $b = 1200s/mm^2$. Additionally, 10 $b = 0$ images were obtained. Scan parameters used were TR=8200ms, TE=100ms, flip-angle $= 90°$, 2.3mm isotropic voxels, matrix size $96 \times 96 \times 61$. To compensate for subject motion and Eddy currents artifacts, the DWI volumes were aligned to the first DWI volume ($b = 0$) using an affine image transformation using normalized mutual information (Collignon et al., 1995). The displacements of the affine transformations were combined with the displacement maps of SPM8's Fieldmap approach (Jezzard and Balaban, 1995), and displacement maps correcting for the non-linearity of the scanner gradients, resulting in a single resampling for each volume, achieved using cubic interpolation. The rotational parts of the affine transformations are extracted and gradient directions corrected using the Finite Strain approach (Alexander et al., 2001).

Tractography
For tractography FMRIB Software Library[2] (FSL) was used. FSL's Bedpostx was used to estimate voxelwise diffusion parameters for each subjects' two DWI sessions. Bedpostx uses Markov Chain Monte Carlo sampling to build distributions of the diffusion parameters and allows for detection of crossing fibers. FSL's Probtrackx (Behrens et al., 2003) was used for probabilistic tractography.

2.3 FCON1000 Resting State fMRI Data

In addition to the DRCMR dataset two other datasets were analyzed. These sets are part of the FCON1000 database and therefore publicly available[3]. We used the Beijing Zang set (part 2), which consists of 42 subjects. Each subject were scanned with 225 volumes, TR=2000ms, and with 33 axial slices. Likewise the Leipzig set was used, which included 37 subjects (21 females), ages=20-42y, each scanned with 195 volumes at a TR=2300ms and with 34 axial slices. The first 5 volumes for these subjects had already been discarded. Using SPM8

---

[2]http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/FSL
[3]http://fcon_1000.projects.nitrc.org/

the remaining volumes were realigned to time series mean and normalized to standard MNI305 space using the template EPI image available in SPM8.

CHAPTER 3

Whole Brain Connectivity

Out of intense complexities intense simplicities emerge.
- Winston Churchill

Section 3.1 gives a short introduction to networks and node clustering. Section 3.2 introduces the Infinite Relational Model, Infinite Hofman-Wiggins, Infinite Diagonal Model, and the Bayesian Community Detection model. Section 3.3 explains the Markov chain Monte Carlo (MCMC) methods for model inference and section 3.4 introduces methods for model evaluations. Section 3.5 summarizes the four papers which use the Bayesian methods for modelling brain networks derived from fMRI and DWI.
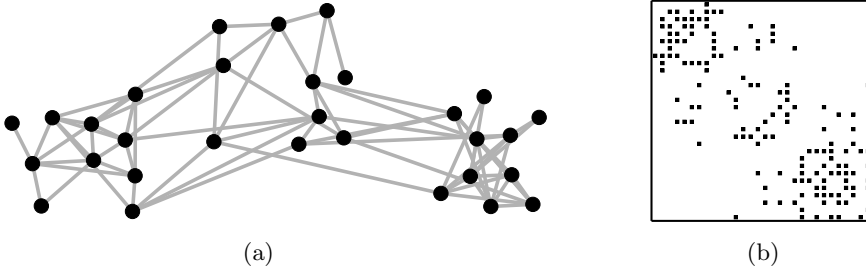
(a)   (b)

Figure 3.1: (a) An undirected and unweighted network. (b) The network's adjacency matrix **A** with dots representing links.

3.1 Network Models

Graph theory provides a mathematical formulation on how nodes in a network interact. It therefore provides a natural framework for many applications wherein the system can be described as a set of nodes and the node interactions can be characterized. Examples of such uses are in modelling social, traffic, citation, and neurobiological systems. A network is a set of nodes, which are connected by links if a relation exists between them. The network is directional if the links have a certain direction and weighted if the links have values different from unity. In the following only simple networks, which are undirected and unweighted networks without loops or multiple edges, are considered. Let J be the number of nodes, then the $[J \times J]$ *adjacency matrix* **A** describes the linking between nodes in the network. For unweighted and undirected networks $A_{ij} = A_{ji} = 1$ if a link between nodes i and j exist, otherwise $A_{ij} = A_{ji} = 0$. The absence of a link is referred to as a nonlink. Figure 3.1 shows an example network along with its adjacency matrix.

3.1.1 Community Structure

Networks can be studied at different levels. At the local level features such as node degree, motifs, clustering coefficient, centrality can be characterized. At the global level features such as link degree distribution and modularity can be quantified.

At the intermediate level networks can be represented in a compressed form by clustering nodes according to how they link to other nodes. A common clustering strategy is to group nodes together such that they have a high internal

link density with low link density to the remaining network. Such clusters are commonly known as *communities*, a term which stems from social networks, where it is observed that persons within a certain community are more related to each other than to persons outside the community. The degree to which a given node partition exhibit community structure can be measured using the *modularity* index (Newman and Girvan, 2004). Informally, the modularity index is defined as the fraction of links which fall within clusters minus the expected number of within-cluster links in a random node degree preserved network. Finding the node partition with optimal modularity is an NP-hard problem (Brandes et al., 2007). However, several methods exist which aims at optimizing modularity, that is, finding the node partition with highest modularity, see e.g., (Lehmann and Hansen, 2007; Newman, 2006).

In neuroimaging the use of graph theory has gained a lot of momentum recently (Bullmore and Bassett, 2011; Sporns, 2011), which is partly due to the increasing number of resting state and diffusion studies conducted. Brain networks are formed by representing brain regions as nodes and either structural or functional connectivity as links. In rs-fMRI the modularity index has been related to motor learning (Bassett et al., 2011), visual working memory (Stevens et al., 2012), normal aging (Meunier et al., 2009), and sleep (Tagliazucchi et al., 2013). Likewise, a number of studies have demonstrated that modularity can be used as bio-marker. E.g., (Alexander-Bloch et al., 2010) showed reduced modularity in patients with childhood-onset schizophrenia and in a recent study Gamboa et al. (2013) showed that patients with early multiple sclerosis had increased modularity compared with healthy controls, an effect which was interpreted as diminished functional integration between functional units. In addition negative correlation between modularity and performance in a dual task was found.

It is, however, not known if modularity reflects the best grouping of nodes. In a modularity optimized node partition the within-cluster link density is optimized without paying attention to the specific linking to other clusters. A better representation at the intermediate level might be formed by partitioning the network into clusters by taking the specific between-cluster linking into consideration. Clusters are herein not only formed by optimizing the internal cluster linking but also by considering how the cluster relates to other clusters. The approach taken in this thesis is to use a family of nonparametric Bayesian models, which are capable of modelling between-cluster link probabilities. By placing different constrains on the probability of observing links between clusters, it can be investigated which model best represent the data. The next section will give a formal description of the nonparametric Bayesian network models considered.

3.2 Nonparametric Bayesian Network Models

This section will introduce the Infinite Relational Model, Infinite Hofman-Wiggins, Infinite Diagonal Model, and Bayesian Community Detection. These models make use of the Chinese Restaurant process, which is also described. The introduction will depart from a description of the Stochastic Blockmodel.

3.2.1 Stochastic Blockmodel

In the Stochastic Blockmodel (Nowicki and Snijders, 2001) the probability of a link between two nodes is determined by cluster membership of the nodes. Let K be the number of clusters and let z_i denote the cluster membership for node i, and let $\mathbf{z} = [z_1, ..., z_J]^\top$ define the cluster membership of all nodes, that is, the node partition. Then the Stochastic Blockmodel defines a matrix $\boldsymbol{\rho}$, with the element $\rho_{z_i z_j}$ being the probability of links between nodes in clusters z_i and z_j. By drawing from the Bernoulli distribution with probability $\rho_{z_i z_j}$ between each pair of nodes i and j, the model can be used to generate networks. The structure of $\boldsymbol{\rho}$ determines the structure of the generated network (Karrer and Newman, 2011). For instance if $\boldsymbol{\rho}$ have zeros in the off-diagonal the generated network will consist of disconnected clusters. If the values in the diagonal of $\boldsymbol{\rho}$ are large compared to off-diagonal values the network will conform with the definition of a community structured network, where nodes within a cluster have dense connections while being sparsely connected to the rest of the network. In the Stochastic Blockmodel the number of clusters K is a fixed finite number and have to be set by the experimenter or learned using model selection.

3.2.2 Chinese Restaurant Process

The nonparametric models described below make use of the Chinese Restaurant Process (CRP) (Aldous, 1985) as prior for the node partition and thereby allows for an infinite number of possible clusters.

In general terms is the CRP a process for generating a partition of a set of N objects. Imagine a Chinese restaurant with infinitely many round tables each with infinite capacity. The first costumer to arrive at this restaurant sits at table1. The second costumer then chooses to sit at table1 with probability $\frac{1}{1+\alpha}$ or at an un-occupied table with probability $\frac{\alpha}{1+\alpha}$. In general, the i-th costumer sits at an un-occupied table with probability $\frac{\alpha}{i-1+\alpha}$ or at table k with probability $\frac{n_k}{i-1+\alpha}$, where n_k is the number of costumers already sitting at table k. After

N costumers have been seated the distribution of the costumers (objects) on the different tables (clusters) defines a partition of the N objects. The process is exchangeable, meaning that the order in which the costumers arrive at the restaurant can be permuted without changing the probability of the resulting partition (Aldous, 1985). α controls the number of clusters generated. In the limit of $\alpha \to \infty$ the probability of starting a new table is 1. When $\alpha = 0$ the probability of starting a new table is 0. Thus, the higher α the more clusters generated. Since the CRP is stochastic each draw from the process will generate a new distribution. The expectation and variance of the number of clusters K can be calculated as (Liu, 1996; Teh, 2010)

$$\mathbb{E}[K|J] \quad = \quad \sum_{i=1}^{J} \frac{\alpha}{i-1+\alpha} \simeq \alpha \log\left(1 + \frac{J}{\alpha}\right) \tag{3.1}$$

$$\mathbb{Var}[K|J] \quad = \quad \sum_{i=1}^{J} \frac{\alpha(i-1)}{(i-1+\alpha)^2} \simeq \alpha \log\left(1 + \frac{J}{\alpha}\right) \tag{3.2}$$

Thus, the number of clusters grows logarithmically with the number of objects. The probability of a given partition $\mathbf{z}$ is

$$P(\mathbf{z}|\alpha) = \frac{\alpha^K (n_k - 1)!(\alpha - 1)!}{(J + \alpha - 1)!} = \frac{\alpha^K \Gamma(\alpha) \prod_k \Gamma(n_k)}{\Gamma(J + \alpha)}, \tag{3.3}$$

where J is the number of objects, n_k is the number of objects assigned to cluster k and K is the number of clusters. Figure 3.2 shows 100 distributions (draws) generated from the CRP using $\alpha = 0.5$, $\alpha = 1$, and $\alpha = 5$ respectively each with $J = 1000$ objects. The distributions are sorted according to the size of the largest cluster. Using $\alpha = 0.5$ relatively few clusters are generated while increasing α also increases (on average) the number of clusters.

3.2.3   Infinite Relational Model

The Infinite Relational Model (IRM) (Kemp et al., 2006; Xu et al., 2006) is a nonparametric generalization of the Stochastic Blockmodel. It uses a hierarchical Bayesian framework by using the CRP as prior for the node partition $\mathbf{z}$. Using the CRP, the model is able to learn the number of clusters from the data, and it potentially allows for as many clusters as nodes in the network. A Beta distribution is used as prior for the link probabilities ρ_{kl}, in this thesis a symmetric Beta distribution ($\text{Beta}(\beta, \beta)$) is considered.

In neuroimaging we are often presented with a number of networks defined on the same set of nodes. Consider for instance networks derived from the
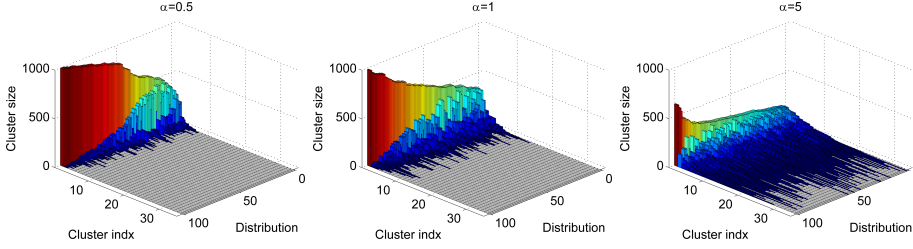
Figure 3.2: 3D histogram showing cluster sizes for 100 distributions generated from the CRP using $\alpha = 0.5$, $\alpha = 1$, and $\alpha = 5$ respectively each distribution having 1000 objects. The distributions have been sorted according to the cluster size of the largest cluster. For $\alpha = 0.5$ relatively few clusters were generated with the largest cluster being relatively large compared to the other clusters. Increasing $\alpha = 1$ increases the probability of generating more clusters.

same brain regions but from different subjects or sessions. The notation $\mathbf{A}^{(n)}$ is used for the n-th network (subject). The IRM can then be defined over all subjects, such that a common node partition is found across all subjects. The link probabilities can either be shared across the group or individual for each subject. For instance Mørup et al. (2010) modelled individual subject link probabilities and in (Andersen et al. (2012), Paper C) and (Andersen et al. (2013), Paper E) the link probabilities were shared across the group of subjects. If not stated otherwise, the multi-network definition of IRM with shared link probabilities will be used.

Thus, the IRM can be summarized as follows.

Infinite Relational Model	
Node partition:	$\mathbf{z} \sim \mathrm{CRP}(\alpha)$
Link probabilities:	$\rho_{kl} \sim \mathrm{Beta}(\beta, \beta)$
Links:	$A_{ij}^{(n)} \sim \mathrm{Bernoulli}(\rho_{z_i z_j})$

The process of generating networks from this model involves first generating a node partition using the CRP. Next, cluster link probabilities are generated from the Beta distribution. Finally, for each network n links are drawn from the Bernoulli distribution according to the link probabilities.

Having defined the model we seek an expression for the conditional posterior distribution of a node's assignment given the assignments of the other nodes. This can then be used to infer the node partition $\mathbf{z}$ using Markov Chain Monte

Carlo (MCMC) methods. The conditional posterior distribution will be derived below while the MCMC procedure will be described in Section 3.3.

For brevity, let the set of N adjacency matrices be defined as $\tilde{\mathbf{A}} = \{\mathbf{A}^{(n)}\}_{n=1}^{N}$. The Bernoulli likelihood for observing $\tilde{\mathbf{A}}$ given $\mathbf{z}$ and $\boldsymbol{\rho}$ is

$$
\begin{aligned}
P(\tilde{\mathbf{A}}|\mathbf{z}, \boldsymbol{\rho}) &= \prod_{n}\prod_{j>i} \rho_{z_i z_j}^{A_{ij}^{(n)}} (1 - \rho_{z_i z_j})^{\left(1 - A_{ij}^{(n)}\right)} \\
&= \prod_{j>i} \rho_{z_i z_j}^{\left(\sum_n A_{ij}^{(n)}\right)} (1 - \rho_{z_i z_j})^{\left(N - \sum_n A_{ij}^{(n)}\right)} \\
&= \prod_{k \geq l} \rho_{kl}^{N_{kl}^{+}} (1 - \rho_{kl})^{N_{kl}^{-}},
\end{aligned}
\tag{3.4}
$$

where N_{kl}^{+} and N_{kl}^{-} are the sum of all links and nonlinks across all networks between cluster k and l, respectively. The Beta probability of $\boldsymbol{\rho}$ is given as

$$
P(\boldsymbol{\rho}|\beta) = \prod_{k \geq l} \frac{\Gamma(2\beta)}{\Gamma(\beta)^2} \rho_{kl}^{\beta-1} (1 - \rho_{kl})^{\beta-1},
\tag{3.5}
$$

where $\Gamma(x) = (x-1)!$ is the Gamma function. By combining Eq. 3.3, Eq. 3.4, and Eq. 3.5 an expression for the joint distribution of $\tilde{\mathbf{A}}$, $\mathbf{z}$, and $\boldsymbol{\rho}$ is

$$
\begin{aligned}
P(\tilde{\mathbf{A}}, \mathbf{z}, \boldsymbol{\rho}|\alpha, \beta) &= P(\tilde{\mathbf{A}}|\mathbf{z}, \boldsymbol{\rho})P(\boldsymbol{\rho}|\beta)P(\mathbf{z}|\alpha) \\
&= \left[\prod_{k \geq l} \rho_{kl}^{N_{kl}^{+}} (1 - \rho_{kl})^{N_{kl}^{-}}\right] \\
&\quad \times \left[\prod_{k \geq l} \frac{\Gamma(2\beta)}{\Gamma(\beta)^2} \rho_{kl}^{\beta-1} (1 - \rho_{kl})^{\beta-1}\right] \\
&\quad \times \left[\frac{\alpha^K \Gamma(\alpha) \prod_k \Gamma(n_k)}{\Gamma(J + \alpha)}\right] \\
&= \left[\prod_{k \geq l} \frac{\Gamma(2\beta)}{\Gamma(\beta)^2} \rho_{kl}^{N_{kl}^{+}+\beta-1} (1 - \rho_{kl})^{N_{kl}^{-}+\beta-1}\right] \\
&\quad \times \left[\frac{\alpha^K \Gamma(\alpha) \prod_k \Gamma(n_k)}{\Gamma(J + \alpha)}\right].
\end{aligned}
\tag{3.6}
$$

Since the Beta distribution is conjugate prior to the Bernoulli distribution the

link probabilities can be integrated out, resulting in

$$
\begin{aligned}
P(\tilde{\mathbf{A}}, \mathbf{z} | \alpha, \beta) &= \int P(\tilde{\mathbf{A}}, \mathbf{z}, \boldsymbol{\rho} | \alpha, \beta) \mathrm{d}\boldsymbol{\rho} \\
&= \left[\prod_{k \geq l} \frac{\mathrm{B}(N_{kl}^+ + \beta, N_{kl}^- + \beta)}{\mathrm{B}(\beta, \beta)} \right] \left[\frac{\alpha^K \Gamma(\alpha) \prod_k \Gamma(n_k)}{\Gamma(J + \alpha)} \right], \quad (3.7)
\end{aligned}
$$

where $\mathrm{B}(x, y) = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}$ is the Beta function. Using Bayes' theorem the conditional posterior distribution of the assignment of a single node z_i is found as

$$
P(z_i = l | \tilde{\mathbf{A}}, \mathbf{z}_{\setminus i}, \beta, \alpha) = \frac{P(\tilde{\mathbf{A}}, \mathbf{z}_{\setminus i}, z_i = l | \alpha, \beta)}{\sum_{l'} P(\tilde{\mathbf{A}}, \mathbf{z}_{\setminus i}, z_i = l' | \alpha, \beta)}, \quad (3.8)
$$

here $\mathbf{z}_{\setminus i}$ is the assignments of all other nodes than node i.

The conditional posterior distribution is used in the MCMC sampling, which will be described in section 3.3.

3.2.4   Infinite Hofman-Wiggins

Hofman and Wiggins (2008) proposed a restricted Stochastic Blockmodel, which only specifies a single within-cluster link probability ρ_w and a single between-cluster link probability ρ_b. Thus, the model restricts all within-cluster link probabilities to be equal and all between-cluster link probabilities to be equal. To be consistent with the other models, the link probability matrix can be formed with the elements

$$
\rho_{kl} = \begin{cases} \rho_w & \text{if } k = l \\ \rho_b & \text{otherwise.} \end{cases} \quad (3.9)
$$

A nonparametric version of this model, termed Infinite Hofman-Wiggins (IHW), was proposed by Mørup et al. (2011). Like in the IRM, a CRP prior is used for the node partition and a symmetric Beta prior is used for the link probabilities. The Bernoulli likelihood is then given as

$$
P(\tilde{\mathbf{A}} | \mathbf{z}, \rho_w, \rho_b) = \rho_w^{N_w^+} (1 - \rho_w)^{N_w^-} \rho_b^{N_b^+} (1 - \rho_b)^{N_b^-}, \quad (3.10)
$$

where N_w^+ and N_w^- are the total number of links and nonlinks which fall within clusters and N_b^+ and N_b^- are the total number of links and nonlinks which fall between clusters. Likewise, the joint Beta probability of ρ_w and ρ_b is given by

$$
P(\rho_w, \rho_b | \beta) = \left[\frac{\Gamma(2\beta)}{\Gamma(\beta)^2} \rho_w^{\beta-1} (1 - \rho_w)^{\beta-1} \right] \left[\frac{\Gamma(2\beta)}{\Gamma(\beta)^2} \rho_b^{\beta-1} (1 - \rho_b)^{\beta-1} \right] (3.11)
$$

Again, the joint likelihood $P(\tilde{\mathbf{A}}, \mathbf{z}, \rho_w, \rho_b | \alpha, \beta)$ can be formed and the link probabilities ρ_w and ρ_b can be integrated out

$$
\begin{aligned}
P(\tilde{\mathbf{A}}, \mathbf{z} | \alpha, \beta) &= \int P(\tilde{\mathbf{A}}, \mathbf{z}, \rho_w, \rho_b | \alpha, \beta) \mathrm{d}(\rho_w, \rho_b) \\
&= \left[\frac{\mathrm{B}(N_b^+ + \beta, N_b^- + \beta) \mathrm{B}(N_w^+ + \beta, N_w^- + \beta)}{\mathrm{B}(\beta, \beta)^2} \right] \\
&\quad \times \left[\frac{\alpha^K \Gamma(\alpha) \prod_k \Gamma(n_k)}{\Gamma(J + \alpha)} \right].
\end{aligned}
\tag{3.12}
$$

Using Eq. 3.8 an expression for the conditional probability of the assignment of a single node given the other nodes can be found and used in the MCMC.

3.2.5 Infinite Diagonal Model

A less restricted version of IRM, termed the Infinite Diagonal Model (IDM), was also proposed in (Mørup et al., 2011). In IDM clusters are allowed different within-cluster link probabilities ρ_{kk}, however, all between-cluster link probabilities ρ_b are shared, forming the link probability matrix

$$
\rho_{kl} = \begin{cases} \rho_{kk} & \text{if } k = l \\ \rho_b & \text{otherwise.} \end{cases}
\tag{3.13}
$$

The Bernoulli likelihood for the IDM is then given as

$$
P(\tilde{\mathbf{A}} | \mathbf{z}, \boldsymbol{\rho}) = \rho_b^{N_b^+} (1 - \rho_b)^{N_b^-} \left[\prod_k \rho_{kk}^{N_{kk}^+} (1 - \rho_{kk})^{N_{kk}^-} \right],
\tag{3.14}
$$

where N_{kk}^+ and N_{kk}^- are the number of links and nonlinks within cluster k and N_b^+ and N_b^- are the total number of links and nonlinks which fall between clusters. The Beta prior for the link probabilities is

$$
P(\boldsymbol{\rho} | \beta) = \frac{\Gamma(2\beta)}{\Gamma(\beta)^2} \rho_b^{\beta - 1} (1 - \rho_b)^{\beta - 1} \left[\prod_k \frac{\Gamma(2\beta)}{\Gamma(\beta)^2} \rho_{kk}^{\beta - 1} (1 - \rho_{kk})^{\beta - 1} \right].
\tag{3.15}
$$

The joint distribution can be formed and $\boldsymbol{\rho}$ can be integrated out

$$
\begin{aligned}
P(\mathbf{A}, \mathbf{z} | \alpha, \beta) &= \int P(\mathbf{A}, \mathbf{z}, \boldsymbol{\rho} | \alpha, \beta) \mathrm{d}\boldsymbol{\rho} \\
&= \left[\frac{\mathrm{B}(N_b^+ + \beta, N_b^- + \beta)}{\mathrm{B}(\beta, \beta)} \right] \left[\prod_k \frac{\mathrm{B}(N_{kk}^+ + \beta, N_{kk}^- + \beta)}{\mathrm{B}(\beta, \beta)} \right] \\
&\quad \times \left[\frac{\alpha^K \Gamma(\alpha) \prod_k \Gamma(n_k)}{\Gamma(J + \alpha)} \right].
\end{aligned}
\tag{3.16}
$$

Again, Eq. 3.8 can be used to derive an expression for the conditional probability of the assignment of a single node given the other nodes.

3.2.6 Bayesian Community Detection

Bayesian Community Detection (BCD) was proposed by Mørup and Schmidt (2012) and is a restricted version of the IRM which conforms to the definition of community structure. The model restricts the between-cluster link probabilities between any two clusters such that it can at most be as high as the lower of the two within-cluster link probabilities. Networks generated with this model will then have lower link density between any two clusters than link density within the clusters. The model is described generatively as follow. First, a node partition is generated using the CRP and the within-cluster link probabilities (diagonal of $\boldsymbol{\rho}$) are drawn from the Beta distribution. Then, for each cluster a cluster gap γ_k is drawn from a Beta distribution and thus takes values between $[0;1]$. The between-cluster link probability between cluster k and l is then generated from an incomplete Beta distribution, BetaInc(a,b,w_{kl}), which is a Beta distribution constrained to the interval $[0;w_{kl}]$. In BCD, w_{kl} is set to the minimum of the cluster gap times the within-cluster link probability of the two clusters, that is, $w_{kl} = \min(\gamma_k \rho_{kk}, \gamma_l \rho_{ll})$. Finally, links are drawn using the Bernoulli distribution according to the cluster link probabilities. The cluster gap controls the degree of community structure. In the limiting case when $\gamma_k = 0, \forall k$ the generated network will be composed of disconnected clusters as there will be no linking between clusters. When $\gamma_k = 1, \forall k$ then the between-cluster link probability can at most take the value of the largest of the two within-cluster link probabilities. In the following the cluster gap is shared across all clusters and is just denoted γ. To summarize the BCD model:

Bayesian Community Detection

Node partition :	$\mathbf{z} \sim \mathrm{CRP}(\alpha)$
Cluster gap :	$\gamma \sim \mathrm{Beta}(v,v)$
Link probability :	$\rho_{kl} \sim \begin{cases} \mathrm{Beta}(\beta,\beta) & \text{if } k = l \\ \mathrm{BetaInc}(\beta,\beta,w_{kl}) & \text{otherwise.} \end{cases}$
	where $w_{kl} = \min[\gamma\rho_{kk}, \gamma\rho_{ll}]$
Links :	$A_{ij}^{(n)} \sim \mathrm{Bernoulli}(\rho_{z_i z_j})$

Let $\dot{\boldsymbol{\rho}} = \{\rho_{kl} | k = l\}$ and $\ddot{\boldsymbol{\rho}} = \{\rho_{kl} | k \neq l\}$ be the sets of within and between cluster link probabilities respectively. Then the joint distribution of $\tilde{\mathbf{A}}, \mathbf{z}, \boldsymbol{\rho}$,

and γ is

$$
\begin{aligned}
P(\tilde{\mathbf{A}}, \mathbf{z}, \boldsymbol{\rho}, \gamma | \alpha, \beta) &= P(\tilde{\mathbf{A}} | \mathbf{z}, \boldsymbol{\rho}) P(\ddot{\boldsymbol{\rho}} | \dot{\boldsymbol{\rho}}, \gamma, \beta) P(\dot{\boldsymbol{\rho}} | \beta) P(\gamma | v) P(\mathbf{z} | \alpha) \\
&= \left[\prod_n \prod_{j>i} \rho_{z_i z_j}^{A_{ij}^{(n)}} (1 - \rho_{z_i z_j})^{1 - A_{ij}^{(n)}} \right] \\
&\quad \times \left[\prod_{k>l} \frac{\rho_{kl}^{\beta-1} (1-\rho_{kl})^{\beta-1}}{\mathrm{B}_{x_{kl}}(\beta, \beta)} \right] \left[\prod_{l=1}^{K} \frac{\rho_{ll}^{\beta-1} (1-\rho_{ll})^{\beta-1}}{\mathrm{B}(\beta, \beta)} \right] \\
&\quad \times \left[\frac{\gamma^{v-1}(1-\gamma)^{v-1}}{\mathrm{B}(v, v)} \right] \left[\frac{\alpha^K \Gamma(\alpha) \prod_k \Gamma(n_k)}{\Gamma(J+\alpha)} \right], \quad (3.17)
\end{aligned}
$$

where B_x is the incomplete Beta function. Integrating over $\ddot{\boldsymbol{\rho}}$

$$
\begin{aligned}
P(\mathbf{A}, \mathbf{z}, \dot{\boldsymbol{\rho}}, \gamma | \alpha, \beta) &= \int P(\mathbf{A}, \mathbf{z}, \boldsymbol{\rho}, \gamma | \alpha, \beta) \mathrm{d}\ddot{\boldsymbol{\rho}} \\
&= \left[\prod_{k=1}^{K} \frac{\rho_{kk}^{N_{kk}^+ + \beta - 1} (1-\rho_{kk})^{N_{kk}^- + \beta - 1}}{\mathrm{B}(\beta, \beta)} \right] \\
&\quad \times \left[\prod_{k>l} \frac{\mathrm{B}_{x_{kl}}(N_{kl}^+ + \beta, N_{kl}^- + \beta)}{\mathrm{B}_{x_{kl}}(\beta, \beta)} \right] \\
&\quad \times \left[\frac{\gamma^{v-1}(1-\gamma)^{v-1}}{\mathrm{B}(v, v)} \right] \left[\frac{\alpha^K \Gamma(\alpha) \prod_k \Gamma(n_k)}{\Gamma(J+\alpha)} \right]. \quad (3.18)
\end{aligned}
$$

Again, using Bayes theorem and eliminating terms which does not depend on ρ_{ll} the marginal posterior reduces to

$$
\begin{aligned}
P(\rho_{ll} | \tilde{\mathbf{A}}, \mathbf{z}, \dot{\boldsymbol{\rho}} \backslash \rho_{ll}, \beta, \alpha, \gamma) &\propto \\
\rho_{ll}^{N_{ll}^+ + \beta - 1} (1-\rho_{ll})^{N_{ll}^- + \beta - 1} &\prod_{k \neq l} \frac{\mathrm{B}_{x_{lk}}(N_{kl}^+ + \beta, N_{kl}^- + \beta)}{\mathrm{B}_{x_{kl}}(\beta, \beta)}. \quad (3.19)
\end{aligned}
$$

The conditional distribution of a node's assignment is given as (Mørup and Schmidt, 2012)

$$
\begin{aligned}
P(z_i = l | \tilde{\mathbf{A}}, \mathbf{z}_{\backslash i}, \dot{\boldsymbol{\rho}}, \beta, \alpha, \gamma) &\propto \\
\rho_{ll}^{r_{il}^+} (1-\rho_{ll})^{r_{il}^-} \alpha^K n_{l \backslash i} &\prod_{k \neq l} \frac{\mathrm{B}_{x_{kl}}(N_{kl}^{+ \backslash i} + r_{ik}^+ + \beta, \ N_{kl}^{- \backslash i} + r_{ik}^- + \beta)}{\mathrm{B}_{x_{kl}}(N_{kl}^{+ \backslash i} + \beta, \ N_{kl}^{- \backslash i} + \beta)}, \quad (3.20)
\end{aligned}
$$

where r_{il}^+ and r_{il}^- are the total number of links and nonlinks from node i to all nodes in cluster l. When terms which does not depend on γ are ignored the posterior reduces to

$$
\begin{aligned}
P(\gamma | \tilde{\mathbf{A}}, \mathbf{z}, \dot{\boldsymbol{\rho}}, \beta, \alpha) &\propto \\
\gamma^{v-1} (1-\gamma)^{v-1} &\prod_{k>l} \frac{\mathrm{B}_{x_{kl}}(N_{kl}^{+ \backslash i} + \beta, N_{kl}^{- \backslash i} + \beta)}{\mathrm{B}_{x_{kl}}(\beta, \beta)}. \quad (3.21)
\end{aligned}
$$

The conditional distributions are used in model inference, which will be described in Section 3.3.

3.2.7   Model Summary

The IRM is the most flexible of the four models and is able to model networks with very complex structure. IRM specifies specific link probabilities between all clusters and is therefore able to cluster a group of nodes with high internal link density into separate clusters if they link differently to other clusters. Likewise, IRM can group nodes with a low internal but high external linking and can therefore capture structure which is beyond the definition of community structure. The BCD also models complex between-cluster linking and has therefore many of the same features as IRM, however, with the constraint that the link probability between any two clusters cannot be greater than either of the two within-cluster probabilities. On the other hand, IDM and IHW only specifies a single between-cluster link probability, which can be seen as a background or noise probability and does not contribute to finding structure in the data.

A nice feature with the nonparametric Bayesian framework is that the experimenter does not have to specifically set the number of clusters. Due to the use of the CRP prior the number of clusters is learned during model inference. The framework also provides an elegant way of inferring the node partition from a group of subjects, with the choice of using individual subject or shared link probabilities.

Figure 3.3 shows a single network generated from each of the 4 models.

3.3   Model Inference

In the previous sections the generative models were defined and expressions for the conditional posterior distributions were derived. This section will describe how the conditional posterior distributions are used in model inference.

3.3.1   IRM, IHW, and IDM Model Inference

In the previous sections the generative models were defined, which allow one to generate data by drawing samples from the model. However, we are interested
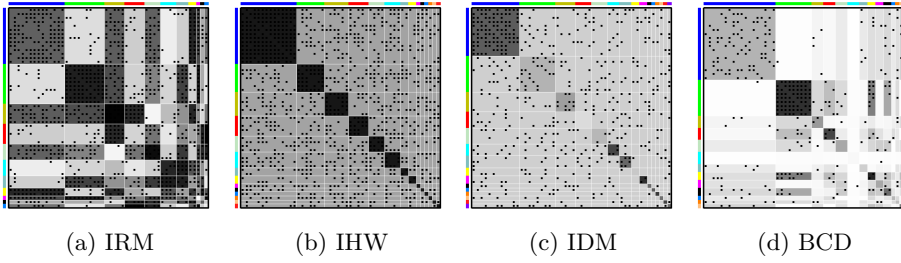
| (a) IRM | (b) IHW | (c) IDM | (d) BCD |

Figure 3.3: Networks generated from each of the four models with $\alpha = 1, \beta = 1$ and for BCD, $v = 1$. Nodes are grouped according to cluster membership. Dots represent links and the gray-shaded background indicates the cluster link probabilities. The IRM model creates complex structured data where all cluster link probabilities are allowed to be different. In the IHW all within-cluster link probabilities are equal and all between-cluster link probabilities are equal. In IDM all between-cluster link probabilities are equal while the within-cluster link probabilities are individual. In BCD the within-cluster link probability for a given cluster is always greater than or equal to the between-cluster link probabilities to the other clusters, however, differences in the different between-cluster link probabilities are allowed.

in inferring the model parameters given the data. By using the model definitions the joint likelihoods were derived and by using Bayes theorem expressions for the conditional posterior distributions were found. It is then possible to sample from these posterior distribution using MCMC.

In IRM, IHW, and IDM the cluster link probabilities were integrated out, which means that model inference only needs to sample over node assignments $\mathbf{z}$. A sampling scheme using Gibbs sampling in combination with split-merge Metropolis-Hastings updates i used (Jain and Neal, 2004; Kemp et al., 2006; Mørup et al., 2010).

In Gibbs sampling, each node's assignment is in turn updated according to the conditional distribution of that node's assignment given the assignments of the remaining nodes. Since Gibbs sampling update the nodes' assignments one at a time, it might be hard to merge two existing clusters or split a single cluster into two separate clusters. Therefore a restricted split-merge scheme is applied in combination with the Gibbs sampling. The split-merge procedure selects two nodes at random. If the two nodes are currently in two separate clusters, then a new partition is proposed, where all nodes from the two clusters are merged into a single cluster. If the two nodes are currently in the same cluster, then a new assignment is proposed, found using restricted Gibbs sampling, which splits the nodes into two clusters. The restricted Gibbs sampling places the two nodes into

two empty clusters and the remaining nodes are assigned to either of the two clusters based on their conditional probability. This state is referred to as the launch state. The launch state is refined by allowing nodes from the two clusters to be re-assigned to either of the two clusters. The resulting configuration is the split-proposal, which will be the new state in the Markov chain if it is accepted with the Metropolis-Hasting acceptance probability.

3.3.2   BCD Model Inference

In IRM, IHW, and IDM it was possible to integrate out the cluster link probabilities. However, due to the dependency between the link probabilities in BCD it is only possible to integrate the between-cluster link probabilities out. The within-cluster link probabilities and the cluster gap are sampled using Metropolis-Hastings. $\mathbf{z}$ is sampled using Gibbs sampling with split-merge moves as described above, however, proposals for within-cluster link probabilities and the cluster gaps are first generated from their prior distribution (Mørup and Schmidt, 2012).

For model inference, software[1] written by Morten Mørup and Mikkel N. Schmidt is used.

3.4   Model Evaluation

This section describes the measures used for evaluating the models.

3.4.1   Mutual Information

The similarity between two samples is calculated using the normalized mutual information (NMI) between their partitions $\mathbf{z}'$ and $\mathbf{z}''$, given as

$$\text{NMI} = \frac{2\text{MI}(\mathbf{z}', \mathbf{z}'')}{\text{MI}(\mathbf{z}', \mathbf{z}') + \text{MI}(\mathbf{z}'', \mathbf{z}'')}, \tag{3.22}$$

---

[1]The IRM, IDM, and IHW software is available from `http://www.imm.dtu.dk/upload/institutter/imm/isp/teaching/02901-2012/morup_exercises.zip` and the BCD software from `http://www2.imm.dtu.dk/pubdb/views/edoc_download.php/6147/zip/imm6147.zip`

where the mutual information (MI) is given as

$$\text{MI}(\mathbf{z}', \mathbf{z}'') = \sum_{k=1}^{K'} \sum_{l=1}^{K''} P(\mathbf{z}' = k, \mathbf{z}'' = l) \log \left(\frac{P(\mathbf{z}' = k, \mathbf{z}'' = l)}{P(\mathbf{z}' = k)P(\mathbf{z}'' = l)} \right), \quad (3.23)$$

where K' and K'' are the number of clusters in $\mathbf{z}'$ and $\mathbf{z}''$, respectively.

3.4.2 Test Log Likelihood

If a model with shared $\boldsymbol{\rho}$ is inferred on a set of networks, the test log likelihood of another set of networks can be calculated. Let $\mathbf{z}'$ and $\boldsymbol{\rho}'$ be the partition and corresponding link probabilities learned from a set of networks. Then the test log likelihood of another set of networks $\tilde{\mathbf{A}}'' = \{\mathbf{A}''^{(n)}\}_{n=1}^{N}$ can be calculated as

$$\log P(\tilde{\mathbf{A}}'' | \boldsymbol{\rho}', \mathbf{z}') =$$
$$\frac{1}{N} \sum_{n=1}^{N} \sum_{j>i} \left[A_{ij}''^{(n)} \log(\rho'_{z'_i z'_j}) + (1 - A_{ij}''^{(n)}) \log(1 - \rho'_{z'_i z'_j}) \right]. \quad (3.24)$$

It is also possible to calculate the test log likelihood for a model with individual $\boldsymbol{\rho}$ if two or more networks are available per subject.

3.4.3 ROC AUC

The Area Under Curve (AUC) of the Receiver Operating Characteristic (ROC) can be calculated on data held out from model inference. Held-out data can either be links and nonlinks, which is treated as missing during inference, or entire networks not included in the inference procedure. AUC is measured by calculating the link probabilities of held out data using the inferred model parameters. By thresholding the probabilities at different levels, the ROC of true positives versus false positives can be made and the area under this curve indicates how well the model predicts test data. AUC = 1 indicates perfect prediction where AUC = $\frac{1}{2}$ is prediction at random.

3.5 Summary of Papers

Three of the included papers in this thesis are using the methods described in the above sections. A fourth paper propose a new model based on the IRM,

which is able to infer shared clustering structure across different network types. This section will give a short summary of the papers.

3.5.1   Paper D, Identification of Functional Clusters in the Striatum Using Infinite Relational Modeling

This paper utilized the IRM for segmentation of the human striatum. The striatum is the main input structure to the human basal ganglia and specific parts of the striatum receives projections from different cortical areas (Haber, 2003). Thus the aim of this paper was to use IRM to cluster voxels in the striatum for identification of these different input zones. The DRCMR rs-fMRI dataset (see section 2.2) was used and voxels of the putamen and caudate nucleus was extracted. Using the rs-fMRI time series from these voxels the correlation matrix was constructed for each of the 30 healthy controls included. The correlation matrices were thresholded to include the highest 5000 positive correlations, which then constitutes the adjacency matrices. We used the IRM model which modelled individual subject link probabilities $\boldsymbol{\rho}^{(n)}$ but had a shared node partition across subjects. Using $\alpha = \beta = 1$, we ran inference with 100 different initializations each for 1000 iterations. The sample with the highest value of the posterior distribution for each run was used as point estimate.

The found node partitions were very consistent across initializations as measured with NMI between each pair of partitions (mean=0.78, std=0.04). From the 100 partitions the co-occurrence matrix was created, which counts how often each node-pair was clustered together. Using this co-occurrence matrix we performed agglomerative hierarchical clustering based on average linkage and thresholded this hierarchy to achieve the final node partition. The resulting clusters were found to represent interhemispheric homologue areas in left and right striatum. The paper also investigated the influence of varying the hyperparameters α and β. Setting $\beta = 1$, we varied $\log_{10}(\alpha)$ from -15 to 15. In this range the MI/NMI between splits and number of found clusters were relatively stable. Setting $\alpha = 1$, we varied $\log_{10}(\beta)$ from -6 to 2, which had a greater impact on both the reproducibility and the number of clusters inferred by the model. The best MI was found with $\beta = 0.1$. MI, however, is dependent on the number of clusters, which can be seen from the high correlation between the curves for MI and number of clusters. The NMI is less dependent on the number of clusters and optimal repetition (NMI=1) was found at $\beta = 100$. However, here only 2 clusters were estimated and thus this solution does not provide much information.

To evaluate the model's ability to predict missing links we treated at random 2.5% of the links and an equivalent number of non-links as missing. The AUC

score was used to evaluate how well the model predicted these missing entries. Across all 30 subjects we found the mean (std) AUC score to be 0.83 (0.06) which is well above chance for all subjects.

In summary, this paper show how the IRM model can be used to segment voxelwise brain regions into functionally similar clusters. Furthermore, the reproducibility of clusterings were evaluated using a split-half framework where MI and NMI were used to evaluate clustering similarity. The model's ability to predict missing links/nonlinks was demonstrated and we found well above chance prediction in all subjects.

3.5.2   Paper C, Identifying Modular Relations in Complex Brain Networks

This paper introduce a general framework for evaluating Bayesian models. The evaluation framework builds on the NPAIRS (nonparametric prediction, activation, influence, and reproducibility resampling) split-half re-sampling (Strother et al., 2002) by estimating the models' predictability and reproducibility using several half-splits of the subject sample. The data sample was split into two equally sized groups and the models were inferred on both splits. The reproducibility was estimated as the NMI between node partitions for the two splits and the predictability as the test likelihood. A model with shared ρ was used, which allow the calculation of test log likelihood.

IRM was evaluated against IDM and IHW using both synthetic data and rs-fMRI data. Three synthetic datasets were generated from each of the three generative models. Each dataset consisted of 20 networks each with 100 nodes. On the data which contains complex structure (IRM generated), the IRM was superior to the two other models both in predictability and reproducibility. Data generated from the two other models does not contain the same relational structure as the off-diagonal elements of the link probability matrices are all equal. In this case the IRM did not overfit the data and produced results comparable with the other two models. In addition to the synthetic data the models were evaluated on rs-fMRI data. Using the Automated Anatomical Labelling (AAL) atlas (Tzourio-Mazoyer et al., 2002) the mean signal in each of the 116 included regions were calculated and the 116×116 correlation matrix were formed for each subject. The correlation matrices were thresholded to include the top 1000 positive correlations forming the adjacency matrices. The models were evaluated using the same framework as on the synthetic data. The results showed that the IRM was superior to the two other models in both reproducibility and predictability, suggesting that rs-fMRI data is better modelled with a model that allows complex interactions between clusters.

To summarize on this paper; using the NPAIRS evaluation framework we showed that the IRM was superior to the other models when data exhibit complex relations. When data were generated from the two simpler models, IRM did not severely overfit and produced comparable predictability and reproducibility as compared with the models which generated the data. IRM fitted the rs-fMRI data best, suggesting that networks derived from whole-brain rs-fMRI exhibit complex relations.

3.5.3   Paper E, Coordinated Activation in the Resting Brain is Community Structured - (*under review*)

In this paper we extended the analysis of rs-fMRI data using Bayesian network modelling. The same framework, as presented in Paper C, was used to evaluate the models. The three models IRM, BCD and IDM were used to model rs-fMRI data from three cohorts. Using the full DRCMR data sample, we ran model inference using the three methods with 10 different initializations. The reproducibility within each method was very high (mean MNI > 0.96 for all methods). Likewise, the clusterings between IRM and BCD were very similar with mean MNI $= 0.94$. In contrast, the clusterings found with IDM was less similar to IRM and BCD, with mean NMI of 0.76 and 0.75, respectively. When using the NPAIRS split-half framework, we showed that IRM and BCD were better than IDM to predict unseen data, which were the case for all three datasets and link densities investigated. For small link densities IDM showed comparable reproducibility compared to IRM and BCD, however, for link densities greater than 8% both IRM and BCD were clearly superior to IDM. Among IRM and BCD the latter showed better or on par reproducibility and predictability for all datasets and link densities.

The two main findings in this paper were that (i) predictability and reproducibility of whole-brain rs-fMRI were greatly improved when modelling the complex relations between clusters (IRM and BCD models) and (ii) since BCD showed greater or on par performance compared to IRM we concluded that the organization of rs-fMRI networks is consistent with a community structured model.

To accompany this paper a toolbox was made[2]. This toolbox is able to construct networks from fMRI data and run the IRM, BCD, IHW, and IDM models as well as make simple visualizations of the results. A screenshot of the GUI is shown in Figure 3.4.

---

[2]Available from `https://brainconnectivity.compute.dtu.dk/`
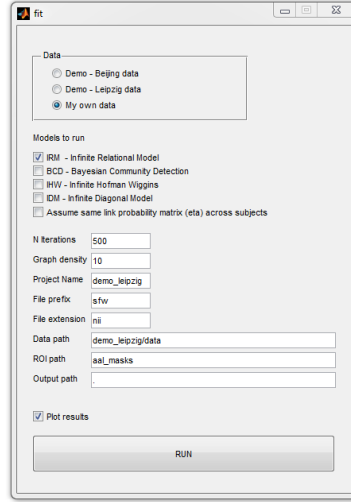
Figure 3.4: Toolbox for generating networks from fMRI data and running the IRM, BCD, IHW, and IDM models. The toolbox is also able to make simple visualization of the results.

3.5.4 Paper F, Joint Modelling of Structural and Functional Brain Networks - (*in preparation*)

Analyzing fMRI and DWI in a joint model allows one to potentially benefit from the complementary information the two modalities provide. There is a great interest in trying to combine these two modalities to reveal how the structural connections are mediating the functional activity (Damoiseaux and Greicius, 2009). For instance, studies have showed that nodes of the default mode network is also structurally connected (Greicius et al., 2009; Teipel et al., 2010). The aim of this paper was to explore clustering similarities in networks derived from fMRI and DWI.

In this paper we used the fMRI and DWI data from the DRCMR dataset. The fMRI networks were derived as in Paper E and Paper C by calculating the correlation between all regions in the AAL atlas. However, the time series were split into two splits (by blocking into 8 blocks and use 4 random blocks per split) and thereby generated two rs-fMRI networks per subject. To generate DWI networks, the AAL atlas were normalized to each subject's native diffusion space. Each region was in turn used as seed and the remaining regions as target in probabilistic tractography. Using FSL's Probtrackx, 5000 streamlines per voxel in the seed region was send off. For each voxel in the seed region R_A the

count of streamlines reaching target region R_B was found and the average of these counts $d_{R_A \rightarrow R_B}$ was calculated. This number will in general not be the same as the count of seeding in R_B and using R_A as target region, therefore the average of $d_{R_A \rightarrow R_B}$ and $d_{R_B \rightarrow R_A}$ was used as element in the connectivity matrix. This matrix was then thresholded to give the DWI network $\mathbf{A}_{\mathrm{DWI}}$. This was done for both of the two DWI datasets per subject resulting in two DWI networks per subject.

Figure 3.5 plots a fMRI and DWI network from a single subject, panel (i) shows the fMRI network and panel (ii) the DWI network. The two networks show clear differences in structure, for instance the fMRI network have dense connections between right and left hemispheres, while most of the DWI connections are within the same hemisphere. Panel (iii) shows the result on running IRM on fMRI alone. Colors indicate cluster membership and for illustration purposes are all nodes within a cluster connected by links (even though they might not be connected in the given network). In addition, the combined adjacency and link probability matrices are shown for both the fMRI and DWI networks. The resulting clusters are in all cases interhemispheric homologue areas. The permuted $\mathbf{A}_{\mathrm{fMRI}}$ shows clear structure with most link densities being high or near zero. Permuting $\mathbf{A}_{\mathrm{DWI}}$ according to the same clustering (found with fMRI data), the link probability matrix is less structured with many elements having intermediate values. Panel (iv) shows the result on running IRM on DWI alone. Here, most clusters are within the same hemisphere and again shows the corresponding permuted adjacency matrix (permuted $\mathbf{A}_{\mathrm{DWI}}$) a more clear structure than the permuted fMRI adjacency matrix (permuted $\mathbf{A}_{\mathrm{fMRI}}$). Panel (v) shows the result on running IRM on both fMRI and DWI networks, thus forcing the same clustering in the two modalities (allowing for individual link probability matrices for all networks). The clusters found here are a mix of the clusters found in the modality specific analysis, some clusters have only nodes within a hemisphere and some clusters have nodes in both hemispheres.

In Paper F a new model is proposed, which is capable of finding shared clusters between two (or more) types of networks. In the general case the model is defined over n network types (modalities) each having m networks (subjects or repetitions). Here, the model is described in terms of fMRI and DWI networks, however, this could easily be extended to include other modalities, such as EEG, MEG, or cortical-thickness networks, as long as the networs are defined on the same set of nodes. The model is an extension of the IRM model and can be described generative as follows. A Bernoulli variable per node determines whether the node is shared across fMRI and DWI. If a node is shared, the cluster in which it participate, is the same for both fMRI and DWI. On the contrary, if the node is individual, it is clustered differently in fMRI and DWI. The cluster membership for the shared nodes are drawn from a common CRP, while cluster memberships for individual nodes are drawn from two separate (fMRI

and DWI specific) CRPs. Each network is associated with its own cluster link probabilities, which are distributed according to a Beta distribution. Finally, for each network (subject) and network type (fMRI and DWI) links are drawn from a Bernoulli distribution with the specific subject-modality-cluster linking probability. The model is schematically shown in Figure 3.6. The blue and red nodes are shared across network type since they are members of the same clusters for both modalities. The remaining nodes are individual (un-shared) across network type since they exhibit different clustering structure in the two modalities.

Using two sets of fMRI and DWI networks per subject, we ran the model within modality (fMRI1 vs. fMRI2, and DWI1 vs. DWI2) and between-modality (fMRI1 vs. DWI1, and fMRI2 vs. DWI2). The analyses were done at the single subject level. By inspecting the fraction of shared nodes in the within-modality analysis, the clustering consistency between scans can be revealed. In the between-modality analysis, the fraction of shared nodes reveals if common structure can be found between modalities. In the between-modality analysis the new model (partially-shared) were compared with a fully-shared model and a fully-unshared model. The fully-shared model assumed that all nodes were shared across network types, which corresponds to running a single IRM. The fully-unshared model assumed that all nodes were individual, which corresponds to running two separate (modality specific) IRM models. The between-modality analysis was evaluated using predictability and reproducibility.

In the within-modality analysis, a high clustering consistency was found between DWI networks with above 90% shared nodes between the two sets. This was seen for all link densities considered. The fMRI networks showed intermediate amount of consistency between sets (about 90% shared nodes at 3% link density which decreased to below 50% shared nodes for 50% link density). When modelling fMRI and DWI jointly, however, a low number of shared nodes were found (below 25% for link densities $\geq 10\%$). The partially-shared model was generally on par with the fully-unshared model, indicating low information shared between modality. Though, for low link densities (3% and 5%) a minor improvement in predicting fMRI links/nonlinks with the fully-shared and partially-shared models was observed. However, this comes with a decrease in predicting DWI networks. The partially-shared model did not improve the reproducibility either.

To summarize, a new nonparametric Bayesian model, which can infer common node clustering across network types, was presented. fMRI and DWI networks showed only limited amount of shared structure. When testing the model within modality, a high consistency was found in the DWI networks and intermediate amount of shared clusters in the fMRI networks. The model is general and can be used for other applications, such as detecting group differences between subject

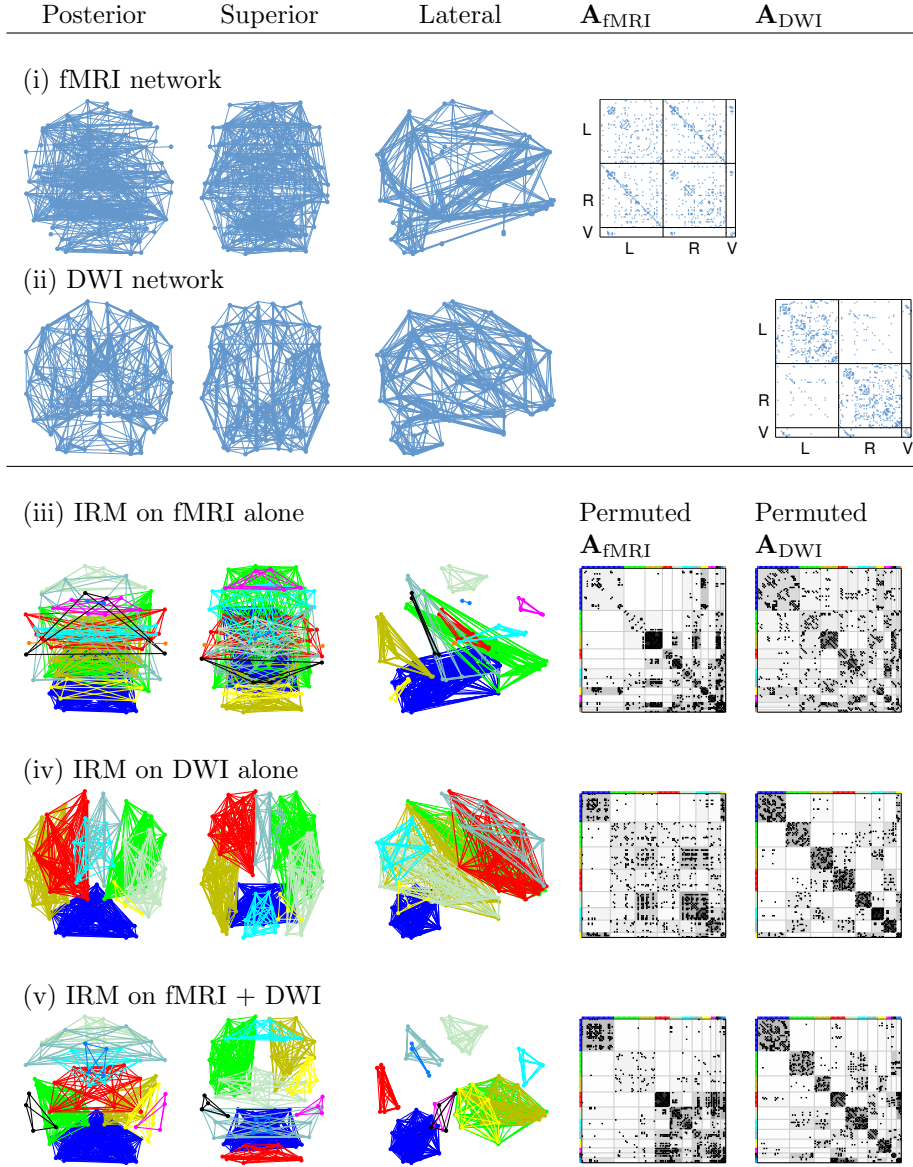populations. This, however, requires further model validation and testing.

Posterior	Superior	Lateral	$\mathbf{A}_{\text{fMRI}}$	$\mathbf{A}_{\text{DWI}}$

(i) fMRI network

(ii) DWI network

(iii) IRM on fMRI alone

Permuted $\mathbf{A}_{\text{fMRI}}$ ... Permuted $\mathbf{A}_{\text{DWI}}$

(iv) IRM on DWI alone

(v) IRM on fMRI + DWI

Figure 3.5: (i) Posterior, superior, and lateral views on a fMRI network from a single subject. (ii) Same views on the DWI network from same subject. (iii) Clustering found when running IRM on fMRI networks. Colors indicate cluster membership and fMRI adjacency matrix ($\mathbf{A}_{\text{fMRI}}$) and DWI adjacency matrix ($\mathbf{A}_{\text{DWI}}$) have been permuted according to IRM clustering found on fMRI alone. (iv) Clustering found when running IRM on DWI alone. The adjacency matrices are permuted according to IRM clustering (found from DWI networks). (v) IRM results found on running IRM on combined fMRI and DWI.
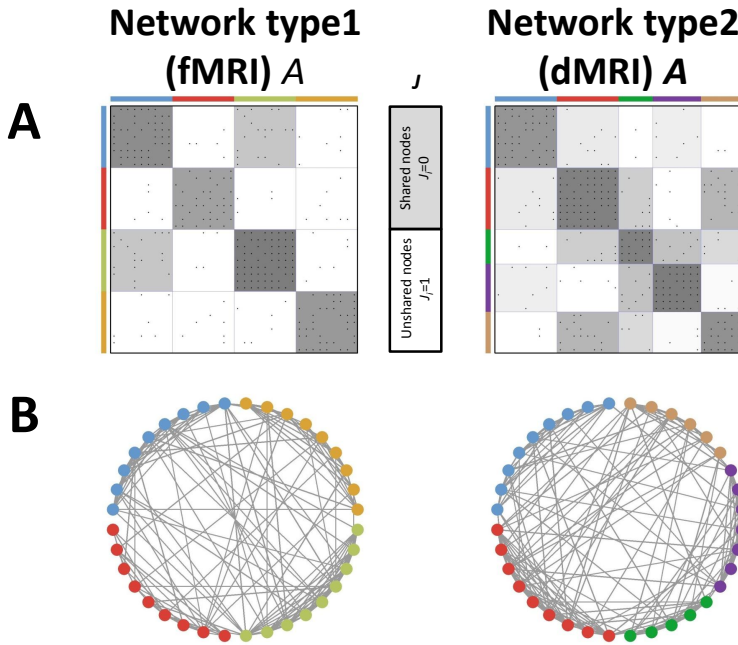
Figure 3.6: Schematic model description. Panel A shows the adjacency matrices for two network types. The small black dots indicate edges. Grouping structures are indicated with colors to the left and above the adjacency matrices and group link probabilities are indicated with the gray-scaled background. In this example nodes belonging to the first two modules are shared across network type while the remaining nodes show different grouping in the two networks. Panel B gives a graphical layout of the networks where nodes are shown with filled circles and edges as lines. The colors indicate the same grouping structure as in panel A.

Local Brain Connectivity

Everything should be made as simple as possible, but not simpler.
- Albert Einstein

This chapter presents work done on investigating how the local homogeneity of rs-fMRI is effected in patients with multiple sclerosis (MS). Contrary to the methods presented in Chapter 3, which considers whole-brain connectivity, this work looks at brain connectivity at the local level. Section 4.1 shortly describes MS. Section 4.2 describes how Kendall's Coefficient of Concordance (KCC) can be used to estimate the local homogeneity of fMRI signals. Section 4.3 summarizes the paper on local homogeneity alterations in MS.

4.1   Multiple Sclerosis

Multiple Sclerosis (MS) is an autoimmune, inflammatory neurological disease (Goldenberg, 2012), which attacks the myelin sheath around axons in the central nervous system. This damage affects the cortico-cortical and cortico-subcortical signal transmission (Compston and Coles, 2002; Lowe et al., 2002; Waxman, 2006). The disease leads to a wide range of disabilities including psychical, cognitive, and mental disabilities. Most MS patients (85-90%) (Lublin and Reingold, 1996) starts in the relapsing-remitting (RR) phase which is characterized by unpredictable events of attacks (relapses) followed by remission periods. 65% of RR-patients will transfer to the secondary-progressive (SP) phase (Confavreux and Vukusic, 2006) where the symptoms will get worse in between attacks. The patient's disability due to MS can be measured using the Expanded Disability Status Scale (EDSS) (Kurtzke, 1983), which is a scale between 0-10, where 0 means no disability and 10 means death due to MS.

4.2   Regional Homogeneity

Alternative to modelling long-range functional brain connectivity the regional homogeneity (ReHo) approach, as the name suggests, models the regional or the local homogeneity of the fMRI signal. ReHo was suggested by Zang et al. (2004) as an alternative method for analyzing rs-fMRI data. Abnormal local signal homogeneity has been found in pathological diseases such as neuromyelitis optica (Liang et al., 2011), Alzheimer's (He et al., 2007), Parkinson's disease (Wu, Long, Zang, Wang, Hallett, Li and Chan, 2009) and Schizophrenia (Liu et al., 2006). Regional homogeneity is typically calculated using Kendall's Coefficient of Concordance (Kendall and Smith, 1939) (KCC) in small $[3 \times 3 \times 3]$ cubic regions of the brain. KCC estimates the agreement of fluctuations in the time series and can be seen as a measure of correlation between a number signals. Calculating KCC between voxels in a small neighborhood then reveals how well synchronized the voxels' time series are in that neighborhood. Since KCC is a rank correlation it does not depend on the absolute amplitude of the signals.

The KCC is calculated as follows. Each of the T time points of a voxel's fMRI signal is given a rank according to signal amplitude. For the N voxels considered the sum of ranks across voxels for timepoint t is denoted r_t. The sum of squares of r_t is then calculated as $s = \sum_{t=1}^{T}(r_t - \bar{r})^2$, where $\bar{r}$ is the mean across all r_t. KCC is then

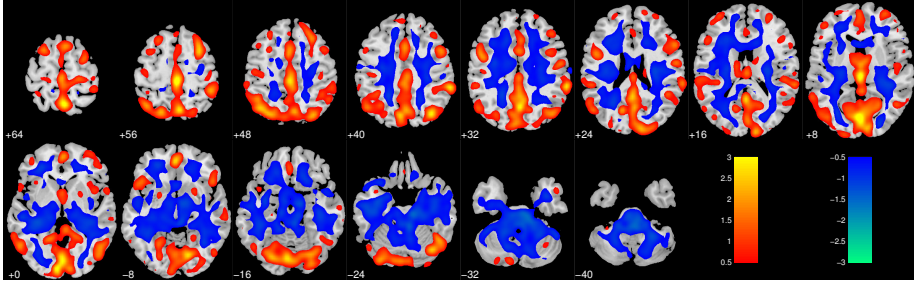$$\text{KCC} = \frac{s}{\frac{1}{12}N^2(N^3 - N)}. \tag{4.1}$$

Figure 4.1: Z-score normalized KCC map for a single subject. Red values indicate areas with increased KCC and blue areas with decreased KCC compared to the mean KCC value in the brain.

The denominator in the equation above is the maximal value of s, which is obtained when all ranks across voxels are equal, and thus ensures that KCC normalizes to 1. The greater KCC the greater (fMRI) signal synchrony.

Normally KCC is calculated in a $[3 \times 3 \times 3]$ cubic region of the brain and the value assigned to the center voxel. This is repeated for the whole brain to generate a KCC-map. Since spatially smoothing of the EPI-images prior to calculating KCC would artificially increase the KCC values, spatial smoothing is applied to the KCC-map. To render the KCC values more normally distributed, the KCC maps are often z-score normalized (Zuo et al., 2010) by subtracting off the mean KCC value of all brain voxels and divide by the standard deviation of KCC values in the same voxels. When normalized, the KCC values represent number of standard deviations away from the mean KCC across the brain.

4.2.1   Relation to Graph Theory

There is a close relation between KCC and the mean of pairwise Spearman rank correlations between all pairs of voxels. Specifically if the Spearman rank correlation between voxels i and j is denoted π_{ij}, then the mean of pairwise Spearman rank correlations is calculated as $\bar{\pi} = \frac{1}{\frac{1}{2}N(N-1)} \sum_{i<j} \pi_{ij}$. Then KCC can be calculated as (Legendre, 2005)

$$\text{KCC} = \frac{(N-1)\bar{\pi} + 1}{n} = \frac{N(N-1)\bar{\pi} + N}{N^2} = \frac{2\sum_{i<j} \pi_{ij} + N}{N^2}, \qquad (4.2)$$

which equals the mean across all elements in π (with ones in the diagonal). Thus, in a weighted network where weights are given by Spearman rank correlations, the (adjusted) mean of all links corresponds to KCC.

4.3   Paper B, Multiple Sclerosis Impairs Regional Functional Connectivity in the Cerebellum

A number of studies have identified inter-regional connectivity alterations in patients with MS (Dogonowski, Siebner, Sørensen, Paulson, Dyrby, Blinkenberg and Madsen, 2013; Dogonowski, Siebner, Sørensen, Wu, Biswal, Paulson, Dyrby, Skimminge, Blinkenberg and Madsen, 2013; Lowe et al., 2002; Roosendaal et al., 2010). The motivation for this paper was to study how the local connectivity is changed in patients with MS.

In (Dogonowski et al., 2014) (Appendix B) we estimated ReHo by calculating KCC maps for each of 42 subjects with MS and 30 healthy controls (HC). After z-score normalizing and spatial smoothing we compared voxelwise KCC maps between MS and HC using a non-parametric two-sample T-test with age as covariate as implemented in FSL Randomise[1]. The significance level was set at $p < 0.05$ at the cluster level. Subjects with MS showed a decrease in regional homogeneity in lobule V and VI of the left cerebellar hemisphere. Likewise, a similar region in the right cerebellar hemisphere showed a trend toward significance. Similarly, a cluster in the left cerebellum, comprising regions of Crus I and Crus II which extended into Crus II and left dentate nucleus, showed negative correlation between KCC and EDSS scores. The direction of the correlation means that patients with higher EDSS scores (more disability) showed less regional homogeneity. Also, a cluster in the right cerebellum showed a trend towards significance.

To investigate the cause of the decreased MS regional homogeneity in cerebellum, we extracted the mean KCC values in the two clusters showing group differences (in left and right cerebellum) from the MS subjects. Likewise, we extracted the mean lesion load (LL) of the cerebellar peduncles and corticospinal tract (CST). Significant correlations (controlling for age) between LL of the left cerebellar peduncles and KCC in left ($p = 0.005$) and right cerebellum ($p = 0.016$) was found. Likewise, when the cerebellar peduncles was segregated into inferior, middle, and superior peduncles we found significant correlations between KCC of left cerebellum and LL in inferior ($p = 0.029$) and middle ($p = 0.012$) but not superior ($p = 0.841$) peduncles. Significant correlation was found between right cerebellum and LL in middle peduncles ($p = 0.042$) but not inferior ($p = 0.122$) and superior (0.072) peduncles. No significant correlations was found with the LL in CST. All the correlations had a negative slope meaning that the greater lesion load the less regional homogeneity.

---

[1]`fsl.fmrib.ox.ac.uk/fsl/fsl-4.1.9/randomise/index.html`

Kernel Principal Component Analysis Denoising

When I have clarified and exhausted a subject, then I turn away from it, in order to go into darkness again.
- Carl Friedrich Gauß

This chapter introduces Kernel Principal Component Analysis (KPCA), which is a non-linear generalization of linear PCA. KPCA can be used for denoising, however, two important parameters, namely the number of components and the scale used for the Gaussian kernel, must be set. Paper A presents a principled way of selecting these two parameters.

(a) Input space   (b) Feature space   (c) Input space
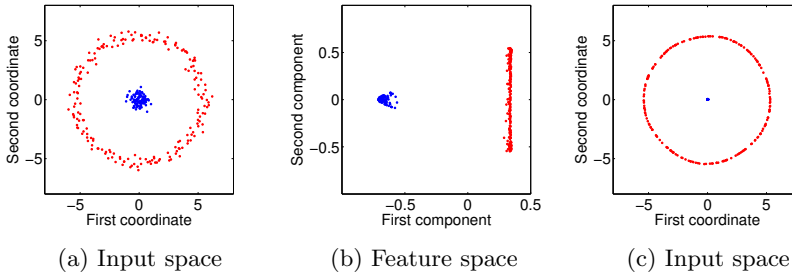
Figure 5.1: Example illustrating the idea behind denoising using KPCA. (a) The 2D data in input space. (b) Data mapped to feature space using the Gaussian kernel with $\sigma^2 = 4$ and projected onto the first two principal components. (c) The denoised data (preimages) in input space. Here data was denoised using the iterative scheme presented in (Mika et al., 1999) using 11 components.

5.1   Kernel Principal Component Analysis

Principal Component Analysis (PCA) is a standard method to do feature extraction, dimensionality reduction, visualization, or denoising. PCA works in the linear domain by projecting data onto the directions with largest variance. However, if data lies on a non-linear manifold linear PCA might not work well. KPCA is a non-linear generalization of PCA. When using KPCA for data denoising the data is projected onto more general non-linear manifolds by mapping data to feature space using a non-linear mapping function. In feature space linear PCA is applied and when data is projected onto a number of the largest eigenvectors and mapped back to input space the so-called pre-image is recovered, where the hope is that noise has been removed from the data.

The process involved in KPCA denoising is illustrated in Figure 5.1. (a) shows the two-dimensional data consisting of blue data points inside a circle of red data points. (b) shows data projected onto the two first principal components in feature space. Here data was mapped with the Gaussian kernel using $\sigma^2 = 4$. The first component captures distance from the center (in input space) and the second component captures radial distance. (c) shows the pre-images after mapping the data back to input space where noise have been removed.

KPCA has been applied in a number of fMRI applications. In (Rasmussen et al., 2012) KPCA was used for fMRI denoising, where improved classification performance after KPCA denoising was demonstrated. Song et al. (2008) also used KPCA for denoising and showed increased sensitivity in activated voxels. Thirion and Faugeras (2003) used KPCA for dynamic fMRI modelling. Sidhu et al. (2012) used KPCA for dimensionality reduction and feature extrac-

tion for classification of ADHD vs. HC. Likewise, Du et al. (2012) used KPCA features for classification of schizophrenia patients vs. HC. In a very recent paper, Song et al. (2014) used KPCA to remove physiological noise in fMRI data. Abrahamsen and Hansen (2011) considered a sparse pre-image reconstruction and show improved brain state decoding accuracy and more reproducible pre-images.

5.1.1   Mathematical Formulation of KPCA

The following introduction to KPCA follows the derivation in (Bishop, 2006). Let the $\mathcal{D} = \{\mathbf{x}_i\}_{i=1}^{N}$ be the set of N data samples in input space $\mathcal{X}$. In linear PCA the sample co-variance matrix of $\mathcal{D}$ is decomposed using an eigen decomposition, and the principal components then correspond to the eigenvectors of the co-variance matrix.

Non-linear PCA can be done by mapping the data to feature space $\mathcal{F}$ using a non-linear map φ. The map $\varphi(\mathbf{x})$ of $\mathbf{x}$ is called the image of $\mathbf{x}$. Then in feature space linear PCA can be conducted on the co-variance matrix of the mapped data points. However, KPCA makes use of the kernel-trick, which states that a kernel function $g(\mathbf{x}_m, \mathbf{x}_n)$ can be defined such that it equals the inner product of their feature space mapped representations

$$g(\mathbf{x}_m, \mathbf{x}_n) = \varphi(\mathbf{x}_m)^\top \varphi(\mathbf{x}_n). \tag{5.1}$$

The goal is then to formulate the PCA in feature space in terms of the kernel function and thus avoid working directly in feature space. For now assume that the data is centered in feature space $\mathcal{F}$. Then the co-variance matrix is given as

$$\mathbf{C} = \frac{1}{N} \sum_{n=1}^{N} \varphi(\mathbf{x}_n)\varphi(\mathbf{x}_n)^\top. \tag{5.2}$$

The eigen decomposition is then

$$\mathbf{C}\mathbf{u}_i = \lambda_i \mathbf{u}_i, \tag{5.3}$$

where $\mathbf{u}_i$ is the eigenvector corresponding to the eigenvalue λ_i. If $\mathbf{C}$ is substituted with the definition given in Eq. (5.2) then

$$\frac{1}{N} \sum_{n=1}^{N} \varphi(\mathbf{x}_i)(\varphi(\mathbf{x}_n)^\top \mathbf{u}_i) = \lambda_i \mathbf{u}_i, \tag{5.4}$$

where by observing that the inner product $(\varphi(\mathbf{x}_n)^\top \mathbf{u}_i)$ is a scalar, we can conclude that $\mathbf{u}_i$ lies in the span of the mapped datapoints $\varphi(\mathbf{x}_n)$, which means

that it can be written as

$$\mathbf{u}_i = \sum_{n=1}^{N} a_{in}\varphi(\mathbf{x}_n). \tag{5.5}$$

Inserting this into Eq. (5.4)

$$\frac{1}{N}\sum_{n=1}^{N}\varphi(\mathbf{x}_n)\varphi(\mathbf{x}_n)^{\top}\sum_{m=1}^{N}a_{im}\varphi(\mathbf{x}_m) = \lambda_i\sum_{n=1}^{N}a_{in}\varphi(\mathbf{x}_n). \tag{5.6}$$

By multiplying both sides with $\varphi(\mathbf{x}_l)^{\top}$ the equation can be expressed using the kernel function $g(\mathbf{x}_n, \mathbf{x}_m) = \varphi(\mathbf{x}_n)^{\top}\varphi(\mathbf{x}_m)$.

$$\frac{1}{N}\sum_{n=1}^{N}g(\mathbf{x}_l, \mathbf{x}_n)\sum_{m=1}^{N}a_{im}g(\mathbf{x}_n, \mathbf{x}_m) = \lambda_i\sum_{n=1}^{N}a_{in}g(\mathbf{x}_l, \mathbf{x}_n). \tag{5.7}$$

This can be written in matrix notation as

$$\mathbf{G}^2\mathbf{a}_i = \lambda_i N\mathbf{G}\mathbf{a}_i, \tag{5.8}$$

where $\mathbf{G}$ is the kernel matrix with $G_{ij} = g(\mathbf{x}_i, \mathbf{x}_j)$. The $\mathbf{a}_i$s can be found by solving the following eigenvalue problem

$$\mathbf{G}\mathbf{a}_i = \lambda_i N\mathbf{a}_i, \tag{5.9}$$

When having solved for the $\mathbf{a}_i$s the solution is normalized by requiring that the eigenvectors $\mathbf{u}_i$ have unit length, which equals to

$$1 = \mathbf{u}_i^{\top}\mathbf{u}_i = \sum_{n=1}^{N}\sum_{m=1}^{N}a_{in}a_{im}\varphi(\mathbf{x}_n)^{\top}\varphi(\mathbf{x}_m) = \mathbf{a}_i^{\top}\mathbf{G}\mathbf{a}_i = \lambda_i N\mathbf{a}_i^{\top}\mathbf{a}_i. \tag{5.10}$$

The derivation above assumed that the mapped data $\varphi(\mathbf{x})$ was centered. This is in general not the case, however, it can be achieved by

$$\tilde{\varphi}(\mathbf{x}) = \varphi(\mathbf{x}) - \bar{\varphi}, \tag{5.11}$$

where $\bar{\varphi} = \frac{1}{N}\sum_{l=1}^{N}\varphi(\mathbf{x}_l)$ is the mean image. This can be written in terms of the kernel function

$$\tilde{g}(\mathbf{x}_n, \mathbf{x}_m) = g(\mathbf{x}_n, \mathbf{x}_m) - \frac{1}{N}\sum_{l=1}^{N}g(\mathbf{x}_l, \mathbf{x}_m) - \frac{1}{N}\sum_{l=1}^{N}g(\mathbf{x}_n, \mathbf{x}_l) + \frac{1}{N^2}\sum_{j=1}^{N}\sum_{l=1}^{N}g(\mathbf{x}_j, \mathbf{x}_l). \tag{5.12}$$

The eigenvectors $\mathbf{u}_i$ are not retrieved per se, however, the projection b_k of a pattern $\mathbf{x}$ onto the k-th eigenvector can be calculated using kernel functions as

$$b_k = \tilde{\varphi}(\mathbf{x})^{\top}\mathbf{u}_i = \sum_{n=1}^{N}a_{in}\tilde{\varphi}(\mathbf{x})^{\top}\tilde{\varphi}(\mathbf{x}_n) = \sum_{n=1}^{N}a_{in}\tilde{g}(\mathbf{x}, \mathbf{x}_n). \tag{5.13}$$

Likewise, an expression for the projection $P_q\varphi(\mathbf{x})$ onto the first q eigenvectors is given as

$$P_q\varphi(\mathbf{x}) = \sum_{i=1}^{q} b_i \mathbf{u}_i + \bar{\varphi} = \sum_{i=1}^{q} b_i \sum_{n=1}^{N} a_{in}\tilde{\varphi}(\mathbf{x}_n) + \bar{\varphi} \qquad (5.14)$$

5.1.2 The Pre-image Problem

The idea with KPCA denoising is to project $P_q\varphi(\mathbf{x})$ back to input space to recover the so-called pre-image, which hopefully is a less noisy version of $\mathbf{x}$. The exact mapping back to input space might not exist. However, an estimate $\hat{\mathbf{x}}$ which minimizes

$$||\varphi(\hat{\mathbf{x}}) - P_q\varphi(\mathbf{x})||^2 = ||\varphi(\hat{\mathbf{x}})||^2 - 2P_q\varphi(\mathbf{x})^\top \varphi(\hat{\mathbf{x}}) + \text{const.}, \qquad (5.15)$$

can be found. Using the Gaussian kernel

$$g(\mathbf{x}, \mathbf{y}) = \exp\left(\frac{-||\mathbf{x} - \mathbf{y}||^2}{2\sigma}\right) \qquad (5.16)$$

Mika et al. (1999) proposed an iterative fixed point algorithm to find $\hat{\mathbf{x}}$ given as

$$\hat{\mathbf{x}}_{t+1} = \frac{\sum_{i=1}^{N} \tilde{\kappa}_i \exp(-||\hat{\mathbf{x}}_t - \mathbf{x}_i||^2/2\sigma^2)\mathbf{x}_i}{\sum_{i=1}^{N} \tilde{\kappa}_i \exp(-||\hat{\mathbf{x}}_t - \mathbf{x}_i||^2/2\sigma^2)}, \qquad (5.17)$$

with $\kappa_i = \sum_{k=1}^{q} b_k a_{ki}$ and $\tilde{\kappa}_i = \kappa_i + \frac{1}{N}(1 - \sum_{j=1}^{N} \kappa_j)$.

5.2 Paper A, Model Selection for Gaussian Kernel PCA Denoising

In (Jørgensen and Hansen, 2012)(Appendix A) we propose kernel Parallel Analysis (kPA) for model selection in Gaussian KPCA denoising. kPA provides a principled way of selecting the scale σ^2 of the Gaussian kernel and the number of components q to retain. The method builds upon parallel analysis (PA) (Beauducel, 2001; Horn, 1965; Ledesma and Valero-Mora, 2007), which is used in linear PCA to estimate the number of components to retain. PA compares the eigenvalues of the co-variance matrix of the original data with eigenvalues of the co-variance matrix of null-data. The null-data is generated by permuting the data within variables a number of times and for each permutation the
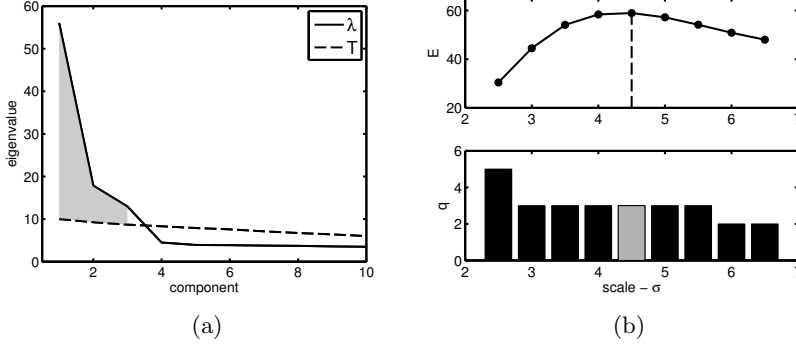
Figure 5.2: Example illustrating kPA. (a) plots the eigenvalues λ of original data (solid line) and the threshold T selected using permuted data (dashed line) for a given σ. The grey area represent the signal energy E. (b) plots the signal energy E (top) and the number of components to retain q (bottom) for different values of σ. Finding the maximum of E selects both the scale σ_{kPA} and the number of components q_{kPA}.

eigenvalues are calculated. For each component a threshold is calculated as the 95th percentile of null-eigenvalues for that component. The number of eigenvalues from the the original data which exceeds this threshold is selected as the dimensionality of the subspace.

kPA takes the same approach as PA, however, the eigenvalues for original and permuted data is calculated in feature space. Thus, this requires that both the original and permuted data are mapped to feature space given a specific σ. In feature space the number components q to retain is found in the same way as in PA, however, note that q will depend on the scale value used, thus we can write $q(\sigma)$. The idea is then to calculate the signal 'energy' as the sum of the retained components' eigenvalues minus the threshold. This signal energy also depends on the scale and kPA therefore chooses the scale σ_{kPA} which maximizes the energy. Once σ_{kPA} is found the number of components follows as $q_{kPA} = q(\sigma_{kPA})$. kPA is illustrated in Figure 5.2 where the eigenvalue spectrum and threshold are plotted and the signal energy corresponds to the gray area. The energy is plotted for different values of σ and the maximum is selected, which then provides an estimate for both σ and q.

kPA is tested using both synthetic data as well as on the USPS dataset of handwritten digits and we show that kPA outperforms other heuristics for selecting the scale as well as dimensionality of the subspace. A Matlab toolbox for performing kPA was developed and is freely available[1].

---

[1] http://www2.imm.dtu.dk/pubdb/views/publication_details.php?id=6236

In a recent report Varon et al. (2013) extended upon the kPA. Especially, the authors note that kPA has difficulties when the dimensionality of data in input space is low. So instead of estimating the eigenspectra from permuted data they estimate the distribution of distances between data points in input space. From this distribution they then generate distance matrices, which are mapped to feature space from which the eigenspectra are computed. They also propose modifications to the fixed-size method as a method to estimate the kernel matrix, which is useful for large datasets. They report improved results especially for low dimensional data using these improvements.

The kPA method is a principled method to determine the two parameters in KPCA. KPCA has already found applications in neuroimaging and future investigations will reveal if kPA is able of finding suitable parameters for denoising in, e.g., fMRI datasets. fMRI data contains noise from a number of sources (scanner, physiological, movement) and a unsupervised method to denoise data would be highly desirable.

CHAPTER 6

Discussion and Conclusion

To arrive at the simplest truth requires years of contemplation.
- Isaac Newton

The main topic of this thesis is network modelling of MR data. The modelling is done at several levels considering both whole brain connectivity as well as modelling of brain connectivity at the local level. In particular, nonparametric Bayesian models were evaluated for use in modelling brain networks. At the local level KCC was used to investigate the local connectivity of fMRI signals and group differences between MS and HC was found. Finally, a principled method to select the parameters for Gaussian KPCA was presented. KPCA is a method with many applications and have already found applications in modelling fMRI data.

At the whole brain level three different nonparametric Bayesian models to infer node clustering were compared. The models differ in the constrains they put on the probability of links between clusters. IRM is the most flexible of the models as this model does not constrain the link probability between clusters. BCD adhere to the definition of community structure and constrains the link probabilities such that a cluster (or community) will have larger within than between cluster linking. Finally, the IDM models individual within-cluster linking but constrains all between-cluster link probabilities to be equal. IDM can be seen as a probabilistic version of the *modularity* index, which is often used for defining communities in brain networks. Among the three models, IRM and BCD were found to be more predictable and reproducible than IDM. This points to the importance of modelling specific between-cluster link probabilities. Comparing the two most expressive models (IRM and BCD) it was found that BCD had higher or on par predictability and reproducibility, with larger difference for the lower link densities. The increased predictability and reproducibility suggests that networks derived from resting state fMRI are community structured.

The model formulation allow one to represent the networks in a compressed form. In particular, a new network can be formed with clusters as nodes and cluster link probabilities as weighted links. These compressed networks can then be used in further analysis, e.g., in group or correlation analysis. This was explored in (Mørup et al., 2010), where the cluster link probabilities were used to classify between MS patients and HC.

Model comparison was achieved with a split-half framework. Modelling shared link probabilities across subjects allows one to evaluate the predictive likelihood for test data. This, together with the reproducibility of clusterings between data splits, forms the evaluation measures. In the model evaluation in Paper C and Paper E a group model was used, which models shared link probabilities across subjects. However, evaluating a model with individual subject link probabilities could be achieved by having multiple datasets per subject. Thereby the subject specific link probabilities can be used for predicting test data for the same subject. This is in fact the approach taken in Paper F where two sets of data were available per subject.

In addition to whole brain network modelling the methods can be used at the local level to, e.g., group voxels into functionally coherent clusters. In Paper D we showed the utility of the IRM model in segmenting voxels in the human striatum. This was achieved by constructing the network by correlating striatal voxels against each other. IRM clustered interhemispheric homologue voxels together indicating that IRM was able to cluster areas involved in same neural computations together. As alternative to segmenting striatum by considering the internal connectivity, a bipartite model could be used. In a bipartite network the connectivities between a set of voxels to another set of voxels are found and forms the basis for analysis. As such, the striatal voxels could be correlated against voxels in cortex. This would result in a clustering of striatal voxels and clustering of cortical voxels and an estimate of the connectivity between elements of the two sets of clusters. In such a model one could test the hypothesis that different areas in striatum exhibit different connectivity to different cortical areas (Di Martino et al., 2008).

Integrating information from structural and functional brain imaging is an area which attracts increasing attention (Damoiseaux and Greicius, 2009; Hinne et al., 2014). Jointly modelling functional and structural connections might reveal information which is not possible to reveal with either modality alone. A novel model based on the IRM, which can infer shared network structure across different network types, is presented. The model allows clusters to be shared or un-shared across network type. The model was tested on networks derived from fMRI and DWI data. Limited amount of shared clustering between these two modalities was found, which might be due to a number of reasons. As was shown in Figure 3.5 there are many between-hemisphere links in the functional networks. However, only a limited amount of streamlines find its way across corpus callosum in the tractography analysis, resulting a lower number of connections between hemispheres for the diffusion network. It has been shown (Honey et al., 2009; Ng et al., 2013) that some of the inconsistencies between functional and structural connectivity can be accounted for by allowing indirect structural connections. In this study we used Pearson correlation as basis for the functional connections. Pearson correlation cannot distinguish direct from indirect connections, that is, if a node A is correlated with node B and B is correlated with C, there is a lower limit on the correlation between A and C (Zalesky et al., 2012). One way to come around this problem, would be to estimate (sparse) partial correlations (Ng et al., 2012) and allow for indirect (second order) structural connections (Bowman et al., 2012). This could render the two types of brain networks more similar and might increase the percentage of shared clustering. When comparing the two DWI networks per subject, we found a high degree of shared clustering between sets, which suggests that even though the networks are derived from probabilistic tractography the two networks are very consistent. Comparing the two rs-fMRI networks per subject, an intermediate amount of sharing was found, which indicates less consistency in resting state

networks. This inconsistency can be due to the non-stationarity of resting state brain correlation, which are reported in, e.g., (Allen et al., 2012). The proposed method is not specific to modelling structural and functional brain networks, but can be used for other types of networks, which are defined on the same set of nodes across network types. An interesting application could be in modelling of group differences in clustering in case-control studies. The model, however, needs further testing before this is possible.

Another approach used in this thesis is to estimate the local functional connectivity by estimating the coherence of the fMRI signals in small neighborhoods. Generating maps reflecting these local connectivities enable us to compare groups of subjects. In particular, we showed that subjects suffering from multiple sclerosis had reduced local connectivity in cerebellum compared to healthy controls, and that the local connectivity decreased with disease progression. This reduced local connectivity in cerebellum is most likely due to lesions in the cerebellar peduncles, as larger lesion load reduced the local connectivity more.

The final topic covered in this thesis in on model selection for denoising using Gaussian KPCA. When denoising using KPCA both the scale of the Gaussian kernel as well as the subspace dimensionality must be selected. To the best of our knowledge our proposed method was the first principled way to simultaneously select both of these parameters. Our method, termed kPA, builds on the Parallel Analysis method for subspace selection in classical PCA. However, in order to also select the Gaussian scale we suggest to optimize the signal energy in feature space, which, when optimized, simultaneously selects both model parameters. KPCA have already found applications in neuroimaging, however, future work will reveal if kPA is able to select suitable parameters when denoising neuroimaging data.

Appendix A

Model Selection for Gaussian Kernel PCA Denoising

Brief Papers

Model Selection for Gaussian Kernel PCA Denoising

Kasper Winther Jørgensen and Lars Kai Hansen

Abstract—**We propose kernel parallel analysis (kPA) for automatic kernel scale and model order selection in Gaussian kernel principal component analysis (KPCA). Parallel analysis is based on a permutation test for covariance and has previously been applied for model order selection in linear PCA, we here augment the procedure to also tune the Gaussian kernel scale of radial basis function based KPCA. We evaluate kPA for denoising of simulated data and the U.S. postal data set of handwritten digits. We find that kPA outperforms other heuristics to choose the model order and kernel scale in terms of signal-to-noise ratio of the denoised data.**

Index Terms—**Denoising, kernel principal component analysis, model selection, parallel analysis.**

I. INTRODUCTION

Kernel principal component analysis (KPCA) is of increasing interest in signal processing, in particular for non-linear signal denoising, see [1] and [2]. While conventional principal component analysis (PCA) denoises signals by projecting onto a linear signal subspace, KPCA denoises by projection onto more general non-linear signal manifolds. A non-linear signal manifold is identified by first mapping the input data to feature space using a non-linear function. In feature space conventional PCA can be applied to extract the main variation in the data by projecting the data onto the subspace spanned by the eigenvectors of the q largest eigenvalues. Finally, the denoised signal is obtained by reconstructing the so-called *pre-image* in input space.

The representer theorem allows effective implementation of the non-linear mapping through inner products represented by the kernel function [3]. Here we will consider the widely used radial basis function aka Gaussian kernel defined by the function $k(x, y) = \exp(-||x - y||^2/2\sigma^2)$. The smoothing scale parameter σ plays an important role to the quality of the pre-image as do the number of principal components retained q. Conventional linear PCA is obtained in the limit $\sigma \to \infty$.

A number of heuristics has been suggested to select the scale. Teixeira *et al*. [1] consider denoising of handwritten digits and they denoise each of the digits in the USPS data set [4] individually and set σ to the maximal distance between each of the training points to the average of all training points. Arias *et al*. [5] set σ as the average distance to the r_σ-th nearest neighbors, $r_\sigma = \{1, 5\}$. Thorstensen *et al*. [6] estimate σ as the median of all mutual distances between all training points. Kwok and Tsang [2] set σ to the mean of all mutual training distances. Likewise, for PCA there exist several methods to estimate the number of components q. The Guttman-Kaiser criterion [7] retains all components with eigenvalues greater than the mean. The so-called Scree criterion plots the eigenvalues in decreasing order and finds the 'elbow' of the eigenvalues spectrum. The lack of a likelihood function for KPCA prevents the use of cross-validation approaches proposed in [8], Alzate and Suykens [9] have proposed an alternative loss function that promotes sparsity, and which also with manual inspection of projection distributions allow model selection. Parallel analysis (PA) [10]–[12] is a resampling based alternative for estimation of q in PCA. PA compares the eigenvalues with the distribution of eigenvalues obtained by PCA on data sets distributed according to a null hypothesis of zero covariance. The PA null distributed data sets are obtained by permuting the measurements among the data points within each feature dimension and q is determined as the set of original PCA eigenvalues greater than the 95th percentile of the corresponding null distribution of eigenvalues.

In this communication, we adapt PA to KPCA to select the model order q and furthermore extend it to automatically select the smoothing scale parameter σ for Gaussian kernels. In particular, we optimize σ to maximize the accumulated eigenvalue advantage of the leading q components compared with PA null data. To our knowledge, this is the first general and automatic scheme for tuning q and σ for KPCA.

II. THEORY

A. KPCA

Let X define the set of N data points $X = [x_1, \ldots, x_N]'$ in input space $\mathcal{X}$. Let φ be a non-linear map from $\mathcal{X}$ to feature space $\mathcal{F}$. The kernel matrix K is constructed from the inner products, i.e., $K_{i,j} = k(x_i, x_j) = \varphi(x_j)'\varphi(x_j)$. The eigen decomposition of the centered kernel matrix is found: $\tilde{K} = HKH = U\Lambda U'$, where $H = I - (1/N)\mathbf{1}\mathbf{1}'$ is the centering matrix, I is the $N \times N$ identity matrix, $\mathbf{1} = [1, 1, \ldots, 1]'$ is a $N \times 1$ vector, $U = [\alpha_1, \ldots, \alpha_N]$ with $\alpha_i = [\alpha_{i1}, \ldots, \alpha_{iN}]'$ is the matrix containing the eigenvectors and $\Lambda = \text{diag}(\lambda_1, \ldots, \lambda_N)$ contains the corresponding eigenvalues [13].

The kth orthonormal eigenvector of the covariance matrix in the feature space can be shown to be

$$v_k = \sum_{i=1}^{N} \frac{\alpha_{ki}}{\sqrt{\lambda_k}} \tilde{\varphi}(x_i) = \frac{1}{\sqrt{\lambda_k}} \tilde{\Phi} \alpha_k \qquad (1)$$

where $\tilde{\varphi}(x) = \varphi(x) - \bar{\varphi}$ is the centered map with $\bar{\varphi} = (1/N) \sum_{i=1}^{N} \varphi(x_i)$ and $\tilde{\Phi} = [\tilde{\varphi}(x_1), \tilde{\varphi}(x_2), \ldots, \tilde{\varphi}(x_N)]$. The projection β_k of the pattern x onto the kth component is then

$$\beta_k = \tilde{\varphi}(x)' v_k = \frac{1}{\sqrt{\lambda_k}} \sum_{i=1}^{N} \alpha_{ki} \tilde{\varphi}(x)' \tilde{\varphi}(x_i)$$
$$= \frac{1}{\sqrt{\lambda_k}} \sum_{i=1}^{N} \alpha_{ki} \tilde{k}(x, x_i) \qquad (2)$$

while the projection $P_q \varphi(x)$ of $\varphi(x)$ onto the subspace spanned by the first q eigenvectors can be found as

$$P_q \varphi(x) = \sum_{i=1}^{q} \beta_i v_i + \bar{\varphi}$$
$$= \sum_{i=1}^{q} \frac{1}{\lambda_i} (\alpha_i' \tilde{k}_x)(\tilde{\varphi} \alpha_i) + \bar{\varphi}$$
$$= \tilde{\Phi} M \tilde{k}_x + \bar{\varphi} \qquad (3)$$

with $M = \sum_{i=1}^{q} (1/\lambda_i) \alpha_i \alpha_i'$ and $\tilde{k}_x = H(k_x - (1/N)K1)$, where $k_x = [k(x, x_1), \ldots, k(x, x_N)]'$.

B. Pre-Image Problem

For denoising we are interested in projecting $P_q \varphi(x)$ back to input space to recover the denoised pattern z - the so-called pre-image. An exact pre-image of $P_q \varphi(x)$ may not exist but a least squares estimate z can be obtained by minimizing

$$||\varphi(z) - P_q \varphi(x)||^2 = ||\varphi(z)||^2 - 2 P_q \varphi(x)' \varphi(z) + \text{const.} \qquad (4)$$

In this brief, we will use the original iterative fixed point algorithm proposed by Mika, Schölkopf *et al.* [3]

$$z_{t+1} = \frac{\sum_{i=1}^{N} \tilde{\gamma}_i \exp\left(-||z_t - x_i||^2 / 2\sigma^2\right) x_i}{\sum_{i=1}^{N} \tilde{\gamma}_i \exp\left(-||z_t - x_i||^2 / 2\sigma^2\right)} \qquad (5)$$

with $\gamma_i = \sum_{k=1}^{q} \beta_k \alpha_{ki}$ and $\tilde{\gamma}_i = \gamma_i + (1/N)(1 - \sum_{j=1}^{N} \gamma_j)$.

C. Kernel Parallel Analysis (kPA)

We extend the idea of PA to KPCA including choice of smoothing scale σ for the Gaussian kernel, and we refer to the proposed method as kPA.

In feature space the eigenvalue λ_i for component i of the PCA is compared with the distribution of eigenvalues of null data sets obtained by permuting the data in input space p times. For component i, the reference threshold T_i is set to the value of the 95th percentile in the component's distribution of eigenvalues. The number of components q to retain is chosen such that the original data eigenvalues are larger than

threshold for all retained components. Note, that both the original data eigenvalues, the reference thresholds, and the number of components q will depend upon the Gaussian scale σ

$$q(\sigma) = \max_{\lambda_i(\sigma) - T_i(\sigma) > 0} i. \qquad (6)$$

A conservative estimate of the signal energy can be obtained as the cumulated difference between the original data eigenvalues and reference threshold levels

$$E(\sigma) = \sum_{i=1}^{q(\sigma)} \lambda_i(\sigma) - T_i(\sigma). \qquad (7)$$

The proposed method chooses the kernel scale σ to *maximize* $E(\sigma)$. The energy is an estimate of the variance of the retained components in kernel space when accounting for the variance of null data. Thus, maximizing the energy in kernel space will maximize the variance of the true signal.

By column-wise permuting the data between samples for a given input dimension, we assure that the null-data is drawn from a distribution which has the same marginal distributions as the original data. Furthermore the input dimensions of the null distribution are statistical independent, i.e., the joint probability density function is fully factorized. This means that all manifold structures in input space are destroyed. Note this is a stronger condition than necessary in PA which only requires a null distribution with no covariance. Hence, the corresponding null distribution in feature space is that of a kernel mapped fully factorized distribution in input space with the correct input space marginals. The kernel spectrum of permuted data represents this "null" information. The distribution of the null kernel spectrum, as estimated by repeated permutation, allows us to determine when structure is present–identified in kPA as eigenvalue magnitudes rejected in the distribution of the null spectrum ($p < 0.05$).

Pseudocode for the kPA algorithm is shown in Algorithm 1. The algorithm starts by making p permutations of the data matrix X. Then the energy is estimated for a number N_σ of different scales σ and the scale σ_{kPA} with maximal energy is chosen with the corresponding number of components $q(\sigma_{kPA})$.

The algorithm calculates the eigenvalues of p kernel matrices. This is repeated for the number of scale values investigated N_σ. The calculation of the data point distance matrix used for the kernel matrix generation can be calculated prior to the iteration over scale values and thus reduces the computations needed. The computational complexity of the eigenvalue calculation is in general $O(N^3)$, where N is the number of data points, though there exist iterative methods for finding the first few eigenvalues of large symmetric matrices which are faster [14]. Thus the worst-case time complexity of the kPA algorithm is $O(p N_\sigma N^3)$.

III. EXPERIMENTS

We use two data sets for illustration. In both data sets, we create noisy data from a set of clean patterns which allows us to measure the quality of the denoising procedure. This experimental design is adapted from the original KPCA

Algorithm 1 kPA

1: Make p permuted replicas of the data matrix by permuting elements in the columns of X independently:
$X_j^{\text{perm}} \leftarrow \text{permute}(X), j = 1, \ldots, p.$

2: Calculate and center the kernel matrix K^{orig} corresponding to the original data matrix X^{orig}:
$\tilde{K}^{\text{orig}} = H K^{\text{orig}} H.$

3: For each permuted dataset calculate and center the kernel matrix K_j^{perm}:
$\tilde{K}_j^{\text{perm}} = H K_j^{\text{perm}} H, j = 1, \ldots, p.$

4: **for** $\sigma = \sigma_{\text{start}} \rightarrow \sigma_{\text{end}}$ **do**

5:   Calculate eigenvalues of kernel matrix:
$\lambda_i(\sigma) \leftarrow \text{eigval}(\tilde{K}^{\text{orig}}, \sigma).$

6:   For the jth permutation calculate eigenvalue for component i in kernel space:
$\tilde{\lambda}_{i,j}(\sigma) \leftarrow \text{eigval}(\tilde{K}_j^{\text{perm}}, \sigma), i = \{1, \ldots, N\}, j = \{1, \ldots, p\}.$

7:   For component i set the threshold T_i to the 95th percentile of eigenvalues of null data:
$T_i \leftarrow$ 95th percentile of $\tilde{\lambda}_{i,*}(\sigma).$

8:   Use (6) to estimate $q(\sigma)$.

9:   Use (7) to estimate $E(\sigma)$.

10: **end for**

11: Select the scale σ_{kPA} which maximizes E.

12: Set the number of components to $q(\sigma_{\text{kPA}})$.

paper [3]. The average signal-to-noise ratio (SNR) over data points was calculated and used as performance metric, where SNR is defined as

$$\text{SNR(dB)} = 10 \log_{10} \frac{\langle S^2 \rangle}{\sigma_{\text{res}}^2} \qquad (8)$$

where S is the noise-free data and σ_{res}^2 is the variance of the residual noise after de-noising.

We used $p = 49$ in the experiments for this brief, which we found resulted in satisfactory results. While increasing p sometimes give more accurate tests this comes with increased computational times.

A. Semi-Circles Simulation

An artificial data set was constructed as two equally populated non-intersecting semi-circles placed initially in a 2-D space ($N = 500$). The two dimensions were both rotated to occupy 25 dimensions generating a $d = 50$ dimensional data set. White Gaussian noise ($\sigma_{\text{noise}} = 0.5$) was added in all 50 dimensions and kPA was used to estimate q and σ before denoising. Fig. 1 shows the eigenvalues of the first ten eigenvectors for the data and the reference threshold level T for null data using $\sigma = 4.5$ and $p = 49$. The shaded area between the two curves where $\lambda_i > T_i$ was next optimized over the single parameter σ. Fig. 2(a) shows the cumulative eigenvalue difference $E(\sigma)$ as a function of the scale value $\sigma \in [2.5, 6.5]$. The maximum of $E(\sigma)$ is attained at $\sigma = 4.5$. Fig. 2(b) shows the number of components $q(\sigma)$ as function of the scale, $q = 3$ components is retained for the optimal σ.
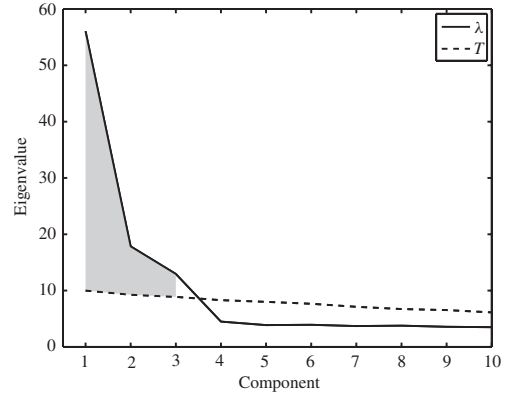

Fig. 1. Simulated data 'semi-circles' ($N = 500$, $d = 50$). The eigenvalues of the empirical data and the null hypothesis 95% percentile reference level. The cumulated eigenvalue difference-'the signal energy' E is the gray area between the two curves. In kPA this area is maximized by tuning the scale of the Gaussian kernel.
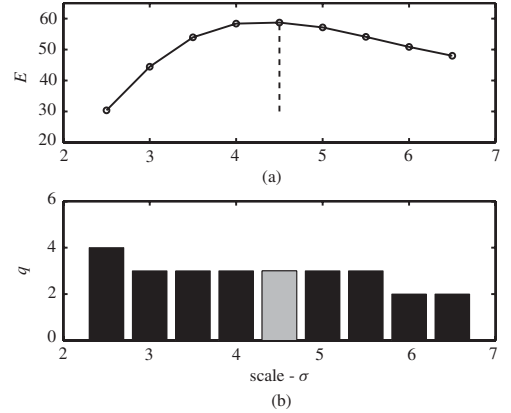

Fig. 2. Simulated data 'semi-circles' ($N = 500$, $d = 50$). (a) $E(\sigma)$ versus kernel scale σ. (b) Corresponding model order q chosen by kPA.

To illustrate the simulation data set we use linear PCA on the noise-free data in input space and project the data onto the two first components obtained on noise-free data. Fig. 3 presents the projected data before noise was added, after noise was added, and after denoising using the optimized parameters found by kPA.

To test kPA more extensively, conditions were varied. The number of data points was varied as $N = \{250, 500, 750\}$ with noise levels $\sigma_{\text{noise}} = \{0.50, 0.75, 1.00\}$. In all cases data were distributed equally between the two semi-circles. kPA was used to estimate σ and q and the data were denoised. Table I shows the estimated q and σ along with SNR mean and standard deviations for 100 repetitions of kPA along with the SNR-optimal combination of the parameters (q, σ) found by exhaustive grid search. For all the nine combinations of sample size and noise level, the kPA estimated parameters remain constant across the 100 repetitions of the experiment.
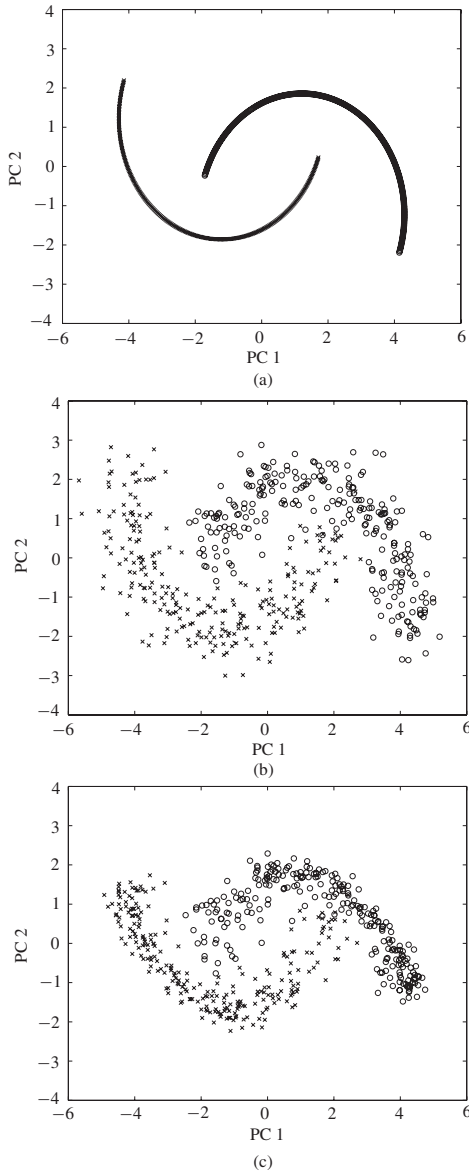
(a)

(b)

(c)

Fig. 3. Simulated data 'semi-circles' ($N = 500, d = 50$). Data in the three panels are all projections on the two first principal components from linear PCA on the noise free in input space. (a) Projections of the noise-free data set onto the two first principal components. (b) PC projections of data with Gaussian white noise added ($\sigma_{\text{noise}} = 0.5$). (c) PC projections of the denoised data using KPCA with $\sigma = 4.5$ and $q = 3$ as determined by kPA.

Fig. 4. Simulated data 'semi-circles'. The SNR (dB) landscape as function of scale parameter σ and the number of components q. Here $N = 500$, $d = 50$, and $\sigma_{\text{noise}} = 0.5$. The kPA solution is indicated by the asterisk at $(q, \sigma) = (3, 4.5)$ and the SNR-optimal solution is indicated with boxes.
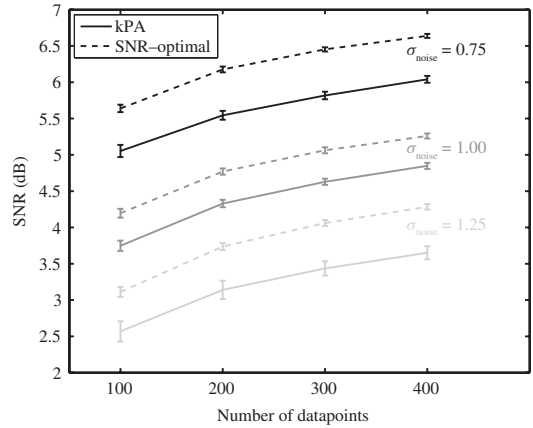
Fig. 5. Mean and standard deviations of SNR (dB) in 100 repetitions of the denoised USPS digits obtained from exhaustive search (SNR-ptimal) and by kPA for three different noise levels.

intermediate value of the two optima, which causes the SNR to drop significantly compared to the SNR-optimal solution.

Fig. 4 shows the SNR-surface when varying q and σ for $(N, \sigma_{\text{noise}}) = (500, 0.5)$. The solution found by kPA $(q, \sigma) = (3, 4.5)$ is indicated with an asterisk while the SNR-optimal solutions are indicated with boxes.

B. USPS Handwritten Digits

kPA was next applied to the USPS database of handwritten digits, often used to illustrate KPCA [3], [4]. Images were normalized to the range [–1;1] and various number of data points $N = \{100, 200, 300, 400\}$ were used equally distributed among the ten digits. Gaussian noise with $\sigma_{\text{noise}} = \{0.75, 1.00, 1.25\}$ was added. We used kPA to determine σ and q and used these parameters to denoise the data and calculate

Likewise, for equal noise level the kPA estimated parameters remain constant across the sample size. Different SNR-optimal parameter combinations were found in the 100 repetitions. For $\sigma_{\text{noise}} = \{0.50, 1.00\}$ the kPA solution is included in the range of SRN-optimal solutions. For $\sigma_{\text{noise}} = 0.75$ the scale takes an
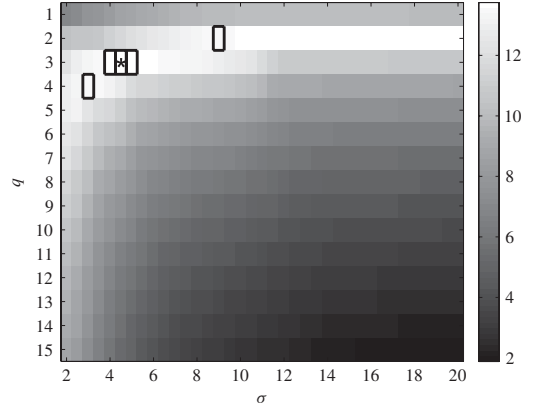
TABLE I

SIMULATED DATA 'SEMI-CIRCLES'. THE ESTIMATED KPCA DIMENSION q, THE SCALE σ, AND MEAN AND STANDARD DEVIATION OF SNR (dB) ESTIMATED WITH KPA VERSUS SNR-OPTIMAL PARAMETERS FOUND BY EXHAUSTIVE GRID SEARCH. THE TEST WAS CARRIED OUT WITH DIFFERENT COMBINATIONS OF DATA SAMPLE SIZE $N = \{250, 500, 750\}$ AND NOISE STANDARD DEVIATION $\sigma_{noise} = \{0.50, 0.75, 1.00\}$. WE REPORT THE PARAMETERS FOUND IN 100 REPETITIONS

		kPA			SNR-optimal		
N	σ_{noise}	q	σ	SNR (dB)	q	σ	SNR (dB)
250	0.50	3	4.5	12.35	2	9.0	12.50
				±0.28	3	4.0, 4.5, 5.0	±0.27
	0.75	2	6.5	8.96	2	8.0, 8.5, 9.0	9.15
				±0.37	3	3.5	±0.37
	1.00	2	8.0	6.44	1, 2	7.0, 7.5, 8.0	6.66
				±0.31		8.5, 9.0	±0.28
500	0.50	3	4.5	13.75	2	9.0	13.77
				±0.23	3	4.0, 4.5, 5.0	±0.22
					4	3.0	
	0.75	2	6.5	10.19	2	8.5, 9.0	10.52
				±0.23	3	3.5, 4.0	±0.23
	1.00	2	8.0	7.89	2	7.0, 7.5, 8.0	7.89
				±0.25		8.5, 9.0	±0.25
750	0.50	3	4.5	14.42	3	4.5, 5.0, 5.5	14.46
				±0.18	4	3.0	±0.18
	0.75	2	6.5	10.71	2	9.0	11.17
				±0.20	3	3.5, 4.0	±0.22
	1.00	2	8.0	8.60	2	7.5, 8.0	8.61
				±0.21		8.5, 9.0	±0.21

TABLE II

USPS DATASET. DENOISING WITH DIFFERENT COMBINATIONS OF NUMBER OF DATA POINTS $N = \{100, 200, 300, 400\}$ AND ADDITIVE NOISE STD. $\sigma_{noise} = \{0.75, 1.00, 1.25\}$. THIS TABLE PRESENTS THE OBSERVED q AND σ IN 100 REPETITIONS OF THE EXPERIMENT USING KPA. IN COMPARISON, THE SNR-OPTIMAL SOLUTIONS OBTAINED FROM EXHAUSTIVE GRID SEARCH ARE SHOWN

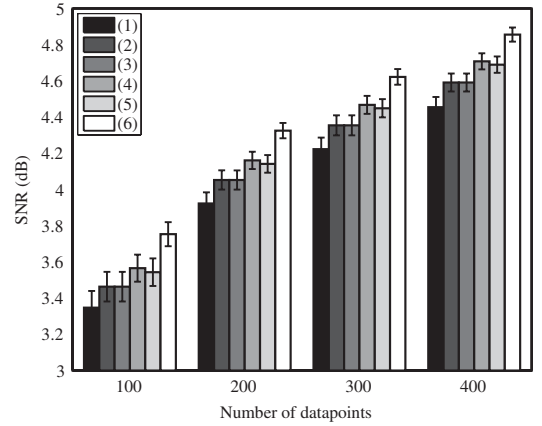|		kPA		SNR-optimal	
N	σ_{noise}	q	σ	q	σ
100	0.75	9–14	15–16	16–21	9–11
	1.00	8–14	19	9–14	10–11
	1.25	6–13	21–23	4–10	11–13
200	0.75	14–18	16	27–32	9–10
	1.00	12–17	19	15–20	10–11
	1.25	9–17	22–23	7–13	11–12
300	0.75	16–20	16	34–42	9
	1.00	14–19	19	18–26	9–10
	1.25	11–19	22–23	9–16	10–13
400	0.75	18–22	16	40–46	9
	1.00	15–20	19	21–28	9–10
	1.25	14–20	22–23	11–18	10–12

Fig. 6. Mean and standard deviations of SNR (dB) in 100 repetitions of the denoised USPS digits obtained by choosing scale according to the six methods. (1) Maximal distance between each training point to average of all training points. (2) Median distance between training points. (3) Mean distance between training points. (4) Average distance to the nearest neighbor. (5) Average distance to the nearest five neighbors. (6) kPA. Here $\sigma_{noise} = 1$, results were similar at other noise-levels. For all data cases kPA significantly outperforms the other five competing methods.

the SNR of the denoised images. Fig. 5 presents the results (mean and standard deviation of 100 repetitions) in terms of the SNR for the kPA solution and the best solution found in exhaustive grid search. The kPA solution is seen to be robust to the number of data points and the noise level used. Table II reports the q and σ values observed in the 100 repetitions under the varying conditions. For both kPA and SNR-optimal solution, the chosen scale value remain constant across N but increases with increasing σ_{noise}. kPA chooses larger scales for all scenarios than the SNR-optimal scale. Both the kPA and SNR-optimal q are increasing with N and decreasing with σ_{noise}. kPAs subspace dimensions q are generally, but not uniformly, smaller than the SNR-optimal solution. The possible tendency to underfit the signal subspace dimension was also noted in [12].

Next, the kPA scale estimate σ was compared with five other heuristics to set the scale: 1) maximal distance between each training point to average of all training points [1]; 2) median distance between training points [6]; 3) mean distance between training points [2]; 4) average distance to the nearest neighbor [5]; and 5) average distance to the nearest five neighbors [5]. For this test, the noise level was set to $\sigma_{noise} = 1.00$ and the number of components was fixed to the q chosen by kPA. Fig. 6 presents the mean and standard

deviation of 100 repetitions for $N = \{100, 200, 300, 400\}$ and shows that kPA outperforms the other methods for all sample sizes investigated, with extremely significant p-values.

The computational complexities for the methods used here is: 1) $O(N)$; 2)–5) $O(N^2)$; and 6) $O(pN_\sigma N^3)$. The mean computational times t_{kPA} for the kPA method in this experiment were $\{N, t_{kPA}\} = \{100, 9.7s; 200, 27.2s; 300, 51.9s; 400, 92.1s\}$, with $N_\sigma = 3$, $p = 49$. For methods
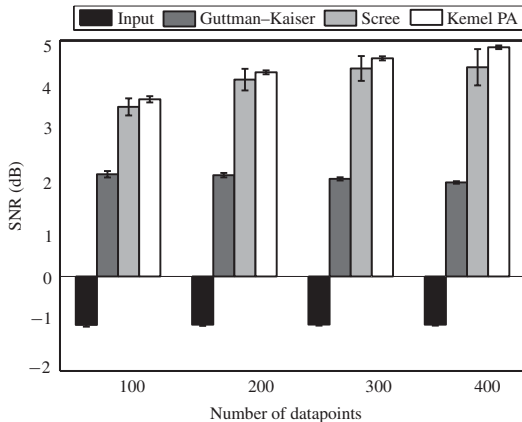
Fig. 7. USPS dataset. SNR (dB) for the input images and the denoised images using the Guttman–Kaiser criterion, the Scree criterion and kPA for choosing the number q of components to retain. The scale σ was chosen by kPA. Here $\sigma_{noise} = 1$, results were similar at other noise-levels. kPA is better than the Scree and clearly superior to the Guttman–Kaiser criterion.

1)–5) the computational times were $t < 0.1$ s for all N. The experiments were done on a Intel(R) Core(TM) i7 CPU, 2.67 GHz system. So, the improved performance of kPA comes with an increased computational time also. This is due to the fact that kPA is based on permutation tests in kernel space while the other methods work on distances in input space.

Finally, we compare methods to choose the number of components q. We compare the kPA solution to the Scree criterion and the Guttman-Kaiser criterion. The Scree criterion was implemented as the first point where the difference between two consecutive eigenvalues in the sorted eigenspectrum was less than 5% of the largest consecutive difference. The Guttman-Kaiser method estimated q as the number of eigenvalues greater than the mean. Fig. 7 plots means and standard deviations of 100 repetitions. kPA significantly outperforms both the Scree and Guttman-Kaiser criteria. In all cases, the differences between kPA and the other methods are extremely significant.

IV. CONCLUSION

We propose kPA, a generalization of PA to KPCA. kPA completes the widely used Gaussian KPCA as an algorithm, as it both solves the subspace dimensionality problem and tunes the smoothing scale parameter. The method optimizes the energy function, which is the accumulated eigenvalue advantage of the leading q components compared with null data. The energy is only a function of the Gaussian kernel smoothing scale, thus the optimization is 1-D. We used two datasets to extensively test the proposed method, namely the artificial semi-circles data and the USPS dataset of handwritten digits. For the semi-circles data, the kPA obtained parameters were shown to be constant across 100 repetitions of the same noise-level and number of data points. Except for $\sigma_{noise} = 0.75$, the chosen parameters are in the range of

SNR-optimal solutions. For $\sigma_{noise} = 0.75$, the kPA solution takes an intermediate value for the scale parameter. For the USPS dataset, we show that the SNR obtained using the kPA solution is robust to the sample size and noise level compared with the SNR-optimal solution. When compared with other heuristics to chose the scale we show that kPA significantly outperform all other methods. Also, when compared to other methods to select the subspace dimensionality the kPA parameter estimates result in significantly higher SNR on the denoised data.

Since kPA is based on permutation tests of the eigenspectra in kernel space, the computational time is larger than the other methods used for comparison in this brief. Future work will focus on improving the computational complexity and test kPA with other noise sources.

REFERENCES

[1] A. Teixeira, A. Tomé, K. Stadlthanner, and E. W. Lang, "KPCA denoising and the pre-image problem revisited," *Digital Signal Process.*, vol. 18, no. 4, pp. 568–580, Jul. 2008.

[2] J. T. Y. Kwok and I. W. H. Tsang, "The pre-image problem in kernel methods," *IEEE Trans. Neural Netw.*, vol. 15, no. 6, pp. 1517–1525, Nov. 2004.

[3] S. Mika, B. Schölkopf, A. Smola, K.-R. Müller, M. Scholz, and G. Rätsch, "Kernel PCA and de-noising in feature spaces," in *Advances in Neural Information Processing Systems*. Cambridge, MA: MIT Press, 1999.

[4] J. J. Hull, "A database for handwritten text recognition research," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 16, no. 5, pp. 550–554, May 1994.

[5] P. Arias, G. Randall, and G. Sapiro, "Connecting the out-of-sample and pre-image problems in kernel methods," in *Proc. IEEE Conf. Comput. Vision Pattern Recognit.*, Minneapolis, MN, Jun. 2007, pp. 1–8.

[6] N. Thorstensen, F. Segonne, and R. Keriven, "Normalization and preimage problem in Gaussian kernel PCA," in *Proc. IEEE Image Process. 15th Int. Conf.*, San Diego, CA, Oct. 2008, pp. 741–744.

[7] K. A. Yeomans and P. A. Golder, "The Guttman-Kaiser criterion as a predictor of the number of common factors," *J. Royal Stat. Soc. Ser. D: Stat.*, vol. 31, no. 3, pp. 221–229, 1982.

[8] L. K. Hansen, J. Larsen, F. A. Nielsen, S. C. Strother, E. Rostrup, R. Savoy, N. Lange, J. Sidtis, C. Svarer, and O. B. Paulson, "Generalizable patterns in neuroimaging: How many principal components?" *NeuroImage*, vol. 9, no. 5, pp. 534–544, May 1999.

[9] C. Alzate and J. Suykens, "Kernel component analysis using an epsiloninsensitive robust loss function," *IEEE Trans. Neural Netw.*, vol. 19, no. 9, pp. 1583–1598, Sep. 2008.

[10] J. L. Horn, "A rationale and test for the number of factors in factor analysis," *Psychometrika*, vol. 30, no. 2, pp. 179–185, 1965.

[11] R. D. Ledesma and P. Valero-Mora, "Determining the number of factors to retain in EFA: An easy-to-use computer program for carrying out parallel analysis," *Pract. Assess. Res. Eval.*, vol. 12, no. 2, pp. 1–11, 2007.

[12] A. Beauducel, "Problems with parallel analysis in data sets with oblique simple structure," *Methods Psychol. Res.*, vol. 6, no. 2, pp. 141–157, 2001.

[13] B. Schölkopf, A. Smola, and K.-R. Müller, "Nonlinear component analysis as a kernel eigenvalue problem," *Neural Comput.*, vol. 10, no. 5, pp. 1299–1319, Jul. 1998.

[14] J. Baglama, D. Calvetti, and L. Reichel, "Iterative methods for the computation of a few eigenvalues of a large symmetric matrix," *BIT Numerical Math.*, vol. 36, no. 3, pp. 400–421, 1996.

Multiple Sclerosis Impairs Regional Functional Connectivity in the Cerebellum

These authors contributed equally to the manuscript.

Multiple sclerosis impairs regional functional connectivity in the cerebellum ☆

Anne-Marie Dogonowski [a,*,1], Kasper Winther Andersen [a,d,1], Kristoffer Hougaard Madsen [a,d], Per Soelberg Sørensen [b], Olaf Bjarne Paulson [a,c], Morten Blinkenberg [b], Hartwig Roman Siebner [a]

[a] Danish Research Centre for Magnetic Resonance, Centre for Functional and Diagnostic Imaging and Research, Copenhagen University Hospital Hvidovre, Kettegaard Allé 30, 2650 Hvidovre, Denmark
[b] Danish Multiple Sclerosis Center, Department of Neurology, Copenhagen University Hospital Rigshospitalet, Blegdamsvej 9, 2100 København Ø, Denmark
[c] Neurobiology Research Unit, Copenhagen University Hospital Rigshospitalet, Juliane Maries Vej 24, Building 9201, 2100 København Ø, Denmark
[d] Cognitive Systems, Department of Applied Mathematics and Computer Science, Technical University of Denmark, Matematiktorvet, Building 321, 2800 Lyngby, Denmark

ARTICLE INFO

ABSTRACT

Resting-state functional magnetic resonance imaging (rs-fMRI) has been used to study changes in long-range functional brain connectivity in multiple sclerosis (MS). Yet little is known about how MS affects functional brain connectivity at the local level. Here we studied 42 patients with MS and 30 matched healthy controls with whole-brain rs-fMRI at 3 T to examine local functional connectivity. Using the Kendall's Coefficient of Concordance, regional homogeneity of blood-oxygen-level-dependent (BOLD)-signal fluctuations was calculated for each voxel and used as a measure of local connectivity. Patients with MS showed a decrease in regional homogeneity in the upper left cerebellar hemisphere in lobules V and VI relative to healthy controls. Similar trend changes in regional homogeneity were present in the right cerebellar hemisphere. The results indicate a disintegration of regional processing in the cerebellum in MS. This might be caused by a functional disruption of cortico-ponto-cerebellar and spino-cerebellar inputs, since patients with higher lesion load in the left cerebellar peduncles showed a stronger reduction in cerebellar homogeneity. In patients, two clusters in the left posterior cerebellum expressed a reduction in regional homogeneity with increasing global disability as reflected by the Expanded Disability Status Scale (EDSS) score or higher ataxia scores. The two clusters were mainly located in Crus I and extended into Crus II and the dentate nucleus but with little spatial overlap. These findings suggest a link between impaired regional integration in the cerebellum and general disability and ataxia.

© 2013 The Authors. Published by Elsevier Inc. All rights reserved.

1. Introduction

Multiple sclerosis (MS) is characterised by disseminated inflammatory demyelination and axonal degeneration in the central nervous system. The disease-related tissue damage delays and disrupts neural signal transmission along cortico-cortical and cortico-subcortical connections (Trapp et al., 1998), causing inefficient information transfer between brain regions. Accordingly, functional magnetic resonance imaging (fMRI) of spontaneous fluctuations in the blood-oxygen-level-dependent (BOLD)-signal during the resting-state has demonstrated changes in long-range functional connectivity of MS patients in the motor network (Dogonowski et al., 2012; Lowe et al., 2002; Roosendaal et al., 2010) and the default-mode network (Hawellek et al., 2011; Rocca et al., 2010).

In addition to studying long-range connectivity within functional brain networks, resting-state fMRI (rs-fMRI) can also be used to assess local connectivity in a brain region. The homogeneity of resting-state BOLD-signal fluctuations in neighbouring voxels reflects local brain connectivity (Zang et al., 2004). Regional homogeneity has successfully been used to identify abnormal local connectivity in pathological conditions such as neuromyelitis optica, Alzheimer's and Parkinson's disease (He et al., 2007; Liang et al., 2011; Wu et al., 2009). Here we employed regional homogeneity analysis of resting-state BOLD-signal fluctuations to test for brain regions where MS patients express an abnormal pattern of local functional connectivity relative to healthy controls. Within the patient group, we were also interested to identify brain regions where the regional expression of local resting-state connectivity would predict clinical disability.

* Corresponding author at: Danish Research Centre for Magnetic Resonance, Copenhagen University Hospital Hvidovre, Department 340, 2650 Hvidovre, Denmark. Tel.: +45 3862 3331; fax: +45 3647 0302.
E-mail addresses: annemd@drcmr.dk (A.-M. Dogonowski), kwjo@dtu.dk (K.W. Andersen), stoffer@drcmr.dk (K.H. Madsen), per.soelberg.soerensen@rh.regionh.dk (P.S. Sørensen), olaf.paulson@nru.dk (O.B. Paulson), blink@dadlnet.dk (M. Blinkenberg), h.siebner@drcmr.dk (H.R. Siebner).
[1] These authors contributed equally to the manuscript.

2. Subjects and methods

2.1. Patients and healthy subjects

Forty-two patients with definite MS fulfilling the revised McDonald criteria (Polman et al., 2005) and 30 sex- and age-matched healthy controls participated in the study. 39 of the 42 patients and 27 of the 30 healthy controls were right-handed as revealed by the Edinburgh Inventory (Table 1) (Oldfield, 1971). Healthy subjects had no history of neurological or psychiatric disease. The study was approved by the scientific ethics committee of Copenhagen and Frederiksberg Communities (protocol no. KF01−131/03 with addendum) and all subjects gave written informed consent.

Patients were recruited from The Danish Multiple Sclerosis Center, Copenhagen, Denmark and comprised 27 patients with relapsing–remitting MS (RR-MS) and 15 patients with secondary progressive MS (SP-MS) (Table 1). Only clinically stable patients who had not experienced a relapse in the three months preceding the magnetic resonance imaging (MRI) measurement were included. Neurological or psychiatric symptoms not attributable to MS were defined as exclusion criteria. Patients were neurologically examined and clinical disability was rated using the Expanded Disability Status Scale (EDSS) (Kurtzke, 1983). The EDSS score ranges from 0 to 10 where 0 equals a normal neurological examination and higher scores indicate more disability. In our patient group, EDSS scores ranged from 0–7 (median score: 4.3). The degree of ataxia was rated using the Multiple Sclerosis Impairment Scale (Ravnborg et al., 1997). The ataxia score evaluated upper and lower limb ataxia and ranges from 0 to 16, where 0 equals no ataxia and the highest scores correspond to the inability to perform coordinated movements. The individual ataxia scores ranged from 0–13 (median score: 3). 81% of the patients (n = 34) presented with ataxia defined as having an ataxia score above 0.

2.2. Magnetic resonance imaging

MRI measurements were performed on a 3 T Magnetom Trio scanner. All rs-fMRI measurements were recorded with a standard single-channel birdcage head-coil using a T2*-weighted echo planar imaging (EPI) sequence with a repetition time (TR) of 2490 ms, an echo time (TE) of 30 ms and a flip angle of 90°. In total 480 whole-brain volumes were acquired over 20 min. A brain volume consisted

of 42 contiguous axial slices with a slice thickness of 3 mm and a 64×64 acquisition matrix covering a field of view = 192×192 mm. The resulting voxel size was $3 \times 3 \times 3$ mm. Subjects were instructed to rest with their eyes closed without falling asleep, and refrain from any voluntary cognitive or motor activity. After the experiment, participants were asked whether they managed to stay awake. All participants reported that they did not fall asleep during scanning. The cardiac cycle was monitored with an infrared pulse oximeter attached to the index finger. Respiration was monitored with a pneumatic thoracic belt. Participants refrained from caffeine, cigarettes or alcohol intake six hours prior to the fMRI-session. Patients continued to take their usual medication.

Additionally, high-resolution three-dimensional structural MRI scans of the brain were acquired using an eight-channel phased array coil (Invivo, FL, USA). A sagittal magnetisation prepared rapid acquisition gradient echo (MPRAGE) sequence (TR = 1550 ms, TE = 3.04, inversion time = 800 ms; flip-angle = 9°) was acquired consisting of 192 contiguous slices with a voxel size of $1 \times 1 \times 1$ mm and an acquisition matrix of 256×256. In addition, sagittal turbo spin echo (TSE) images (TR = 3000 ms, TE = 354 ms) and fluid-attenuated inversion recovery (FLAIR) images (TR = 6000 ms, TE = 353 ms) were obtained. The TSE and FLAIR images covered the whole brain consisting of 192 slices with a voxel size of $1.1 \times 1.1 \times 1.1$ mm and a 256×256 acquisition matrix. The structural scans were used to estimate total lesion load of cerebral white-matter in MS patients and to exclude subclinical white-matter lesions in healthy controls. Whole-brain lesion load was quantified using a semi-automatic lesion segmentation method guided by expert knowledge as described previously (Dogonowski et al., 2012).

Patients with MS showed reduced regional homogeneity in the cerebellum. This prompted us to perform a follow-up analysis on the patient data in which we tested for a linear relationship between lesion load in the cerebellar peduncles and the cortico-spinal tract (CST) and the change in regional homogeneity in the cerebellum. Lesion load in the cerebellar peduncles was estimated by extracting the cerebellar peduncles including the inferior, middle, and superior cerebellar peduncles as defined in the JHU white-matter tractography atlas and split into a left and right region-of-interest (ROI) (Hua et al., 2008). Lesion load of the inferior, middle, and superior cerebellar peduncles was also estimated individually. The ROI applied to estimate lesion load of the left, right and combined CST was specified as defined in JHU white-matter tractography atlas (Hua et al., 2008).

Table 1
Demographics and clinical characteristics.

	Healthy controls	Patients with MS	RR-MS	SP-MS
Number of subjects	30	42	27	15
(men; women)	(15; 15)	(20; 22)	(10; 17)	(10; 5)
Median age in years	45	45	39	51
(range)	(22–69)	(25–64)	(25–59)	(30–64)
Handedness	27; 2; 1	39; 3; 0	26; 1; 0	13; 2; 0
right; left; ambidextrous				
Median disease duration in years	n.a.	11.5	9	20
(range)		(3–43)	(3–27)	(6–43)
Median EDSS score	n.a.	4.3	3.5	6.0
(range)		(0–7)	(0–6.5)	(3.5–7.0)
Median ataxia score	n.a.	3.0	2.0	4.0
(range)		(0–13)	(0–13)	(1–10)
Median whole-brain lesion load in ml	n.a.	21.4 (n = 41)	17.4 (n = 26)	35.8 (n = 15)
(range)		(1.8–126.3)	(1.8–96.6)	(3.7–126.3)
Median lesion load in left and right cerebellar peduncles in ml	n.a.	0.04	0.05	0.04
(range)		(0–1.1)	(0–1.1)	(0–0.8)
Treatment	n.a.	35	24	11
IFN-β	n.a.	6	6	0
Glatiramer acetate	n.a.	9	9	0
Natalizumab	n.a.	5	5	0
Immunosuppresive agents	n.a.	10	5	5
Other	n.a.	6	0	6

MS = multiple sclerosis; RR-MS = relapsing–remitting multiple sclerosis; SP-MS = secondary progressive multiple sclerosis; n.a. = not applicable; EDSS = Expanded Disability Status Scale; IFN = Interferone.

2.3. Resting-state fMRI data analysis

2.3.1. Pre-processing

Image pre-processing of the EPI images used statistical parametric mapping (SPM) software (SPM8, Wellcome Trust Centre for Neuroimaging, http://www.fil.ion.ucl.ac.uk/spm) implemented in Matlab 7.9 (MathWorks, Massachusetts, USA). The first five brain volumes were discarded to account for T1 equilibrium effects. The remaining 475 brain volumes were first realigned to the first volume in the time series, then a mean volume was generated, and the realignment procedure was repeated with this volume as the target. The volumes were then co-registered to the same-session T1-weighted MPRAGE data set using a six-parameter rigid-body transformation. The MPRAGE volumes were spatially normalised to the Montreal Neurological Institute (MNI) 305 standard template using the unified segmentation and normalisation procedure as implemented in SPM8 and the same normalisation parameters were used to normalise the EPI images (Ashburner and Friston, 2005). Spatial smoothing was not done at this step.

2.3.2. Noise filtering

Cardiac and respiratory activity is known to produce signal changes in fMRI time-series (Lund et al., 2006) which could give rise to correlations resembling those observed in rs-fMRI (Birn et al., 2006). In the current study we address these issues by comprehensive filtering of cardiac and respiratory effects (using separate recordings of cardiac and respiratory cycles). Hence, possible confounding effects of residual movement and physiological signals were reduced by filtering several nuisance signals prior to further analysis. The filter was based on a Volterra expansion of the estimated movement parameters obtained during the realignment procedure (24 parameters), a Fourier expansion of the aliased cardiac (10 parameters) and respiratory (6 parameters) cycles and corresponding respiration by cardiac cycle interaction (4 parameters) (Friston et al., 1996; Glover et al., 2000). Also changes in respiration volume over time have shown to produce correlation patterns resembling those observed in rs-fMRI (Birn et al., 2006). To take these changes into account the filter also included 41 delayed versions (from −20 to 20 s) of the respiration volume (Birn et al., 2006). In addition, time-series from white-matter (left and right superior longitudinal fasciculus) and the lateral ventricles were included in the filter. In addition, low frequency scanner drifts were removed by including a discrete cosine basis set with a cut-off frequency of 1/128 Hz.

2.3.3. Regional homogeneity

The regional homogeneity was assessed using the Kendall's Coefficient of Concordance (KCC) measure (also known as Kendall's W). We calculated the KCC for a cube consisting of $3 \times 3 \times 3$ voxels and the KCC-value was assigned to the centre voxel. KCC was calculated as follows: For each voxel each time point was assigned a rank (an integer between 1 and 475) according to signal intensity. The sum of ranks for a given time point i is denoted R_i and the average value of R_i is denoted R_m. The KCC is then calculated as,

$$ \text{KCC} = \frac{\sum_{i=1}^{n} (R_i)^2 - n(R_m)^2}{\frac{1}{12} K^2 \left(n^3 - n \right)} $$

where $K = 27$ is the number of voxels used in the KCC calculation, and $n = 475$ is the number of time points. The KCC calculation was repeated for each voxel in the brain to create a regional homogeneity map (KCC-map). The KCC-map was then z-transformed by subtracting off the mean KCC of the whole brain and dividing by the standard deviation of all KCC-values in the brain mask (Zuo et al., 2010). When transforming to z-scores, the KCC-map represents values relative to the mean KCC in the brain for each individual subject. KCC-maps were created for each subject using in-house software. Finally, the KCC-maps were smoothed with a $6 \times 6 \times 6$ mm full-width at half maximum Gaussian kernel.

2.3.4. Statistical inference

Since the KCC-maps in general are not guaranteed to be normal distributed, statistical inference at group level employed the non-parametric permutation method using 5000 permutations as implemented in the FSL software tool Randomise (http://fsl.fmrib.ox.ac.uk/fsl/randomise/). We corrected for multiple comparisons by controlling the family-wise error (FWE) rate at the cluster level considering all brain voxels. Cluster extent threshold was set at $p < 0.01$ (two-tailed test) and a corrected p_{FWE}-value < 0.05 at the cluster level was considered statistically significant. Clusters with a corrected p_{FWE}-value < 0.15 are reported as trends. The statistical maps were superimposed on a mean normalised T1-weighted image based on the MPRAGE brain scans of all subjects using Mango (http://ric.uthscsa.edu/mango/). Anatomical locations were inferred using the automated anatomical labeling (AAL) atlas (Tzourio-Mazoyer et al., 2002) for the cerebrum and the probabilistic atlas of the human cerebellum (Diedrichsen et al., 2009) to locate the cerebellar clusters.

We performed two main analyses. The aim of the first analysis was to identify between-group differences in the spatial expression of regional homogeneity including all voxels in the brain. To this end, we compared the KCC-maps of MS patients and healthy controls using a non-parametric two-sample permutation test. The second analysis focused on MS patients exclusively and addressed the question where in the brain regional homogeneity of BOLD-signal fluctuations reflects global disease-related disability. Using only the data of MS patients, we tested for a linear relationship between the voxelwise KCC-values and individual EDSS scores using a non-parametric multiple linear regression model. In both analyses, age was treated as a confounding covariate.

Our main analysis revealed that MS patients show an abnormal expression of regional homogeneity in the cerebellum. This finding motivated a set of explorative post-hoc analyses to gain a deeper pathophysiological understanding of the main results. We performed the same between-group analysis again, but we only included the 34 MS patients who clinically displayed signs of ataxia. We also conducted two exploratory analyses comparing the KCC-maps of patients with ataxia versus patients without ataxia and the KCC-maps of patients with RR-MS versus SP-MS using non-parametric two-sample permutation test with age as covariate. We also tested for a linear relationship between the voxelwise KCC-values and individual ataxia scores in the patient group using a similar multiple regression model with age as covariate.

Since the between-group and regression analyses exclusively revealed changes in the cerebellum, one obvious question was whether the altered regional homogeneity in the cerebellum is related to cerebellar lesion load. We estimated regional lesion load of the cerebellar peduncles because they are the input and output structures of the cerebellum. CST lesion load was also estimated because of its importance for sensory-motor functioning and served as a control tract. Lesion load was extracted for right and left cerebellar peduncles and right and left CST. We tested whether inter-individual variations in lesion load of right and left cerebellar peduncles or right and left CST was linearly correlated with a decrease in mean KCC-values in the right and left cerebellar clusters that had been revealed by the between-group comparison while controlling for the effect of age. In each patient, we extracted the KCC-values from all voxels which belonged to the left cerebellar cluster reaching statistical significance in the between-group comparison. The voxels contributing to the mean KCC of the right cerebellar hemisphere belonged to the cluster showing a trend decrease in KCC in MS relative to healthy controls. The correlations were calculated using Matlab's partial correlation controlling for age. Permutation tests were conducted to assess the significance level (10.000 permutations). The regression analyses were repeated including whole-brain lesion

load as covariate. We tested for a correlation between lesion load of each cerebellar peduncle (i.e., bilateral ROIs for the inferior, middle, and superior peduncle) and the CST (i.e., bilateral ROI including right and left CST) with the mean KCC of the cluster in right and left cerebellum controlling for age.

As a few of the subjects included were left-handed, we repeated the original between-group and regression analysis with EDSS scores using a similar model including both handedness and age as covariates. Finally, we tested whether a shorter resting-state fMRI session would yield the same results as the analysis based on the entire data set recorded over 20 min. This analysis was performed because most resting-state fMRI protocols are shorter. We split the resting-state fMRI data in half and repeated the between-group and the regression analysis with EDSS score on the first half of the resting-state fMRI data (238 volumes). In the additional post-hoc analysis, we applied an uncorrected $p < 0.05$ as statistical threshold.

3. Results

3.1. Reduced regional homogeneity in the cerebellum

Patients with MS showed reduced regional homogeneity in the left cerebellar hemisphere relative to healthy controls ($p_{FWE} = 0.020$; peak Z-score = 4.63; peak coordinate (x,y,z): −18, −46, −26). The area displaying a decrease in regional homogeneity was located in the left superior cerebellar lobe, including mainly lobules V and VI extending into lobule IV and the vermis (Fig. 1). Corresponding regions in the right cerebellar lobule VI with extensions into lobule IV and V showed a similar decrease in regional homogeneity, but reached only trend significance ($p_{FWE} = 0.127$; peak Z-score = 4.21 at voxel (x,y,z): 27, −67, −23). A subcortical cluster comprising right caudate nucleus, thalamus and globus pallidus also reached trend significance ($p_{FWE} = 0.123$; peak Z-score = 4.01 at voxel (x,y,z): 18, 2, 22). In contrast, no area in the hemispheric cortex displayed a significant difference in regional homogeneity in patients with MS relative to healthy controls, even at a more liberal statistical threshold ($p < 0.15$, FWE-corrected). The reduced regional homogeneity in left and right superior

cerebellar lobe in MS patients relative to controls could be replicated when removing the patients without ataxia (n = 8) from the analysis ($p_{FWE} = 0.004$; peak Z-score = 4.47 at voxel (x,y,z): 27, −67, −23). The only difference was that the left and right cerebellar clusters merged into one big cluster with the peak voxel localised in the right cerebellar cluster. The main finding, a disease-related reduction in regional homogeneity in patients in the left cerebellar cluster remained significant after including handedness as covariate ($p_{FWE} = 0.022$). The comparison contrasting regional homogeneity in the RR-MS and SP-MS groups yielded no significant differences in regional homogeneity between the phenogroups. Neither were significant differences in regional homogeneity detected between ataxic patients and patients without ataxia. Repeating the main analysis on half of the data, the observed group differences remained significant.

3.2. Relation between regional homogeneity and global clinical disability

Regression analysis identified a single cluster in the left cerebellar posterior lobe where regional homogeneity decreased with increasing clinical disability as expressed by the EDSS score. The more the MS patients were clinically affected (i.e. the higher the individual EDSS score), the weaker was regional homogeneity in the left cerebellar cluster (Fig. 2; $p_{FWE} = 0.031$; peak Z-score = 4.52 at voxel (x,y,z): −21, −82, −32). The cerebellar cluster covered Crus I and extended into Crus II and the left dentate nucleus. A homologous area in the right cerebellum comprising Crus I and parts of Crus II showed a similar relationship between regional homogeneity and individual EDSS scores, but this trend did not reach statistical significance ($p_{FWE} = 0.096$; peak Z-score = 3.98 at voxel (x,y,z): 27, −70, −38). The observed inverse linear relationship between KCC-values and EDSS scores remained significant after including handedness as covariate ($p_{FWE} = 0.033$). A regression analysis that only used the first half of the resting-state fMRI data replicated the linear relation between EDSS score and KCC in the left cerebellar cluster. Additionally, the right cerebellar cluster became significant that failed to reach significance in the original analysis using the entire fMRI data set.
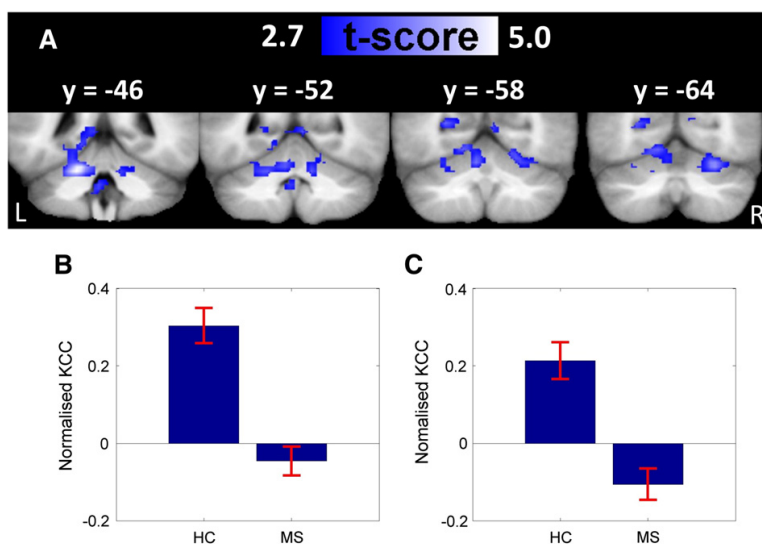

Fig. 1. Decrease in regional homogeneity of cerebellar regions in MS. Coronal t-maps representing voxels with reduced regional homogeneity in MS compared with controls (A). Left (B) and right (C) cerebellar cluster mean normalised KCC and standard error bars of healthy controls and MS patients. KCC = Kendall's Coefficient of Concordance; HC = healthy controls; MS = multiple sclerosis.
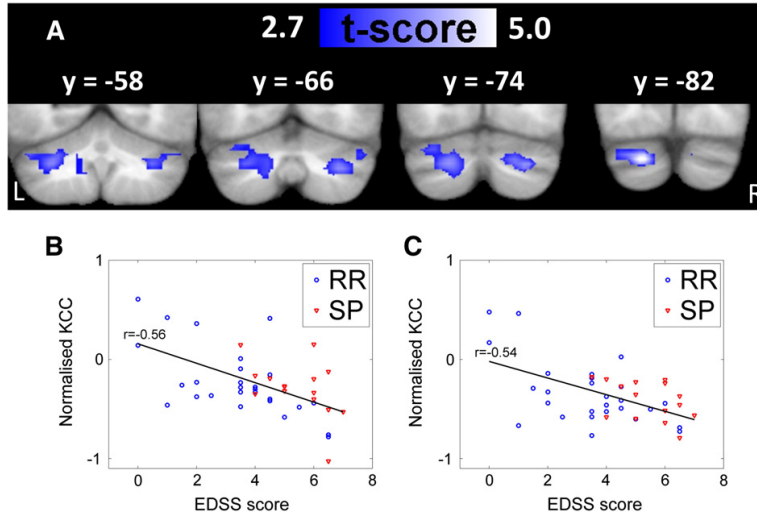
Fig. 2. Regional homogeneity of cerebellar regions correlates with disability in MS. Coronal t-maps representing voxels where KCC correlated with EDSS scores in MS (A). Left (B) and right (C) cerebellar cluster mean normalised KCC for each subject (y-axis) plotted against EDSS scores (x-axis) with a regression line. KCC = Kendall's Coefficient of Concordance; EDSS = Expanded Disability Status Scale; RR = relapsing–remitting multiple sclerosis; SP = secondary progressive multiple sclerosis.

Exploratory post-hoc regression analysis identified a single brain cluster in the left cerebellar posterior lobe where regional homogeneity decreased with increasing ataxia scores (Fig. 3; $p_{FWE} = 0.047$; peak Z-score = 3.61 at voxel (x,y,z): -6, -52, -41). This cluster also comprised Crus II and Crus I, extending into the left dentate nucleus and lobule IX. The cluster showing reductions in regional homogeneity with increasing EDSS scores showed only limited overlap (20% of voxels) with the cluster displaying decreases in regional homogeneity with increasing ataxia scores (Fig. 3), although individual ataxia and EDSS scores were highly correlated (correlation coefficient = 0.68; $p < 0.001$).

3.3. Relation between regional homogeneity and lesion load in the cerebellum

In the 41 patients in whom structural MRI data of sufficient quality were available lesion load of the cerebellar peduncles and CST were extracted. The volume of the right cerebellar peduncle template was estimated to 9.6 ml and the volume of left cerebellar peduncle template 9.9 ml. The maximal lesion load of the cerebellar peduncles accounted for 1.1 ml, corresponding to 5.9% of the total cerebellar peduncles volume (Table 1).

The correlations between cerebellar regional homogeneity and lesion load of the cerebellar peduncles and CST are illustrated in Fig. 4. Lesion load of the left cerebellar peduncle showed a significant partial correlation with the regional mean KCC in the left ($p = 0.005$) and right cerebellum ($p = 0.016$), whereas lesion load of right cerebellar peduncles showed no significant linear relation with right or left cerebellar regional mean KCC ($p \geq 0.097$). The results were reproduced when repeating the regression analyses controlling for whole-brain lesion load. We also segregated lesion load of the bilateral superior, middle, and inferior cerebellar peduncles and correlated these lesion measures with left and right cerebellar regional mean KCC. Lesion load of middle cerebellar peduncles showed a significant partial correlation with left ($p = 0.012$) and right cerebellar regional mean KCC ($p = 0.042$). Lesion load of inferior cerebellar peduncles showed a significant partial correlation with left regional mean KCC ($p = 0.029$), and lesion load of superior cerebellar peduncle showed a trend towards a partial correlation with right cerebellar regional mean KCC ($p = 0.072$). Lesion load in left, right or bilateral CST showed no linear correlation with right or left cerebellar regional mean KCC ($p > 0.2$). The overall lesion load of the right ($p = 0.012$) and left ($p = 0.044$) cerebellar peduncles showed a significant positive relationship with individual EDSS scores.

4. Discussion

Using regional homogeneity of resting-state fluctuations in the BOLD-signal as index of local functional connectivity, we show that local functional connectivity is impaired in the left cerebellum in MS. Regional homogeneity was found to be reduced in lobules V and VI of the left cerebellar hemisphere in patients with MS relative to controls. Further, patients with higher EDSS scores displayed less local connectivity in Crus I and dentate nucleus of left cerebellum. Similar trends were present in homologous regions of the right cerebellum. Post-hoc analyses showed that patients with higher lesion load of the left cerebellar peduncle showed more reduced local cerebellar connectivity.

4.1. Cerebellar regional homogeneity in MS

The impairment of regional functional homogeneity in MS extends previous reports showing marked pathological changes in cerebellar
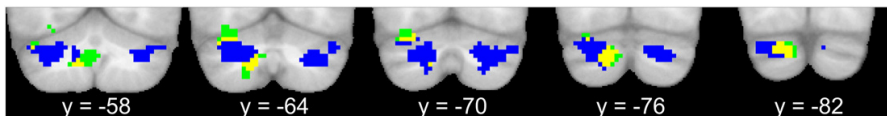

Fig. 3. Regional homogeneity of cerebellar regions correlates with ataxia in MS. Coronal maps representing voxels where KCC correlated with individual ataxia scores in MS (green cluster). The blue cluster represents voxels where KCC correlated with EDSS scores in MS. The significant clusters of the two analyses showed limited spatial overlap (20% of voxels; yellow cluster represented in the foreground) and mainly include Crus I (40%), Crus II (14%), the dentate nucleus (9%), lobule IX (9%), and lobule VI (10%).
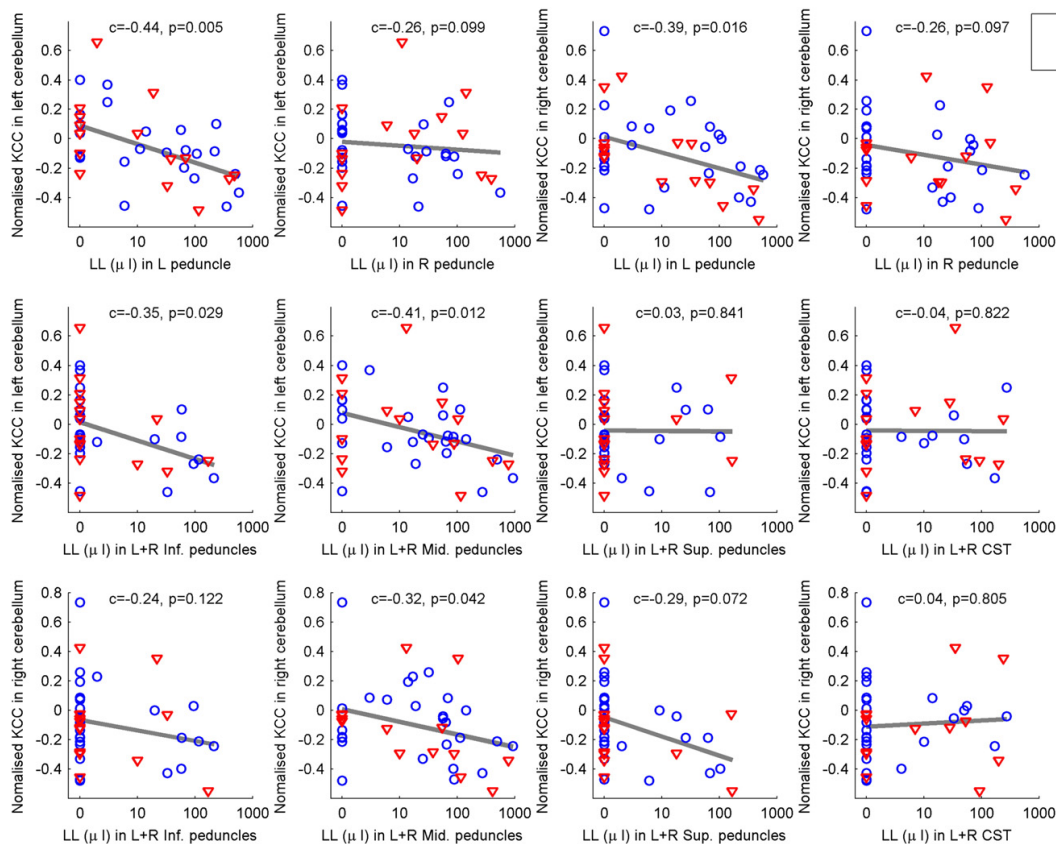
Fig. 4. Correlation between cerebellar regional homogeneity and lesion load of the cerebellar peduncles and cortico-spinal tract. The first row illustrates the correlations between regional homogeneity in left or right cerebellar region-of-interest (ROI) with lesion load of left or right cerebellar peduncles. The second row illustrates the correlations between regional homogeneity of the left cerebellar ROI and lesion load of the bilateral inferior, middle, and superior cerebellar peduncles and cortico-spinal tract (CST). The third row illustrates the correlations between regional homogeneity of the right cerebellar ROI and lesion load of the bilateral inferior, middle, and superior cerebellar peduncles and CST. KCC = Kendall's Coefficient of Concordance; LL = lesion load; L = left; R = right; RR = relapsing–remitting MS; SP = secondary progressive MS; CST = cortico-spinal tract; Inf. peduncles = inferior cerebellar peduncles; Mid. peduncles = middle cerebellar peduncles; Sup. peduncles = superior cerebellar peduncles.

cortex in MS (Anderson et al., 2009; Gilmore et al., 2009; Kutzelnigg et al., 2007). The cerebellum is critically involved in the temporal and spatial integration of neural inputs from both descending cortical afferents (cortico-ponto-cerebellar projections) and ascending spinal afferents (proprioceptive information via spino-cerebellar projections) via the cerebellar peduncles (Ramnani, 2006). The reduced cerebellar regional homogeneity might indicate that the temporo-spatial integration of converging cortico-ponto-cerebellar and spino-cerebellar inputs is disintegrated in patients with MS. The notion of dysfunctional neural integration in the cerebellum is supported by a series of positron emission tomography studies which found bilateral reductions in cerebellar resting-state glucose metabolism in early RR-MS which was ascribed to the remote effect of demyelinating lesions (i.e. crossed cerebellar diaschisis) (Blinkenberg et al., 2000; Derache et al., 2006). In addition, two activation fMRI studies have reported decreased inter-regional functional and effective connectivity during motor task involving cerebellar regions in MS (Rocca et al., 2009; Saini et al., 2004).

Disrupted local connectivity in the cerebellum might be caused primarily by local cortical damage. Alternatively, MS lesions in cerebellar white-matter tracts might cause a functional disruption of neurotransmission in these pathways. The latter hypothesis is supported by the finding that patients displayed more reduced right and left cerebellar regional homogeneity with increasing lesion load of the left cerebellar peduncle. In particular, lesion load in the middle cerebellar peduncle correlated with the decrease in regional cerebellar connectivity. The middle cerebellar peduncle is a main input structure containing axons which project from the nuclei pontis of the opposite side to the cerebellar cortex. This structure–function relationship suggests that the disturbed regional integration of incoming cortico- and spino-cerebellar information increases with lesion load in the cerebellar peduncles, but this hypothesis needs to be tested specifically in future studies.

When interpreting the present results, one needs to take into account the neuroanatomical organisation of the cerebellum. Research has provided converging evidence that Crus I and Crus II represent the 'cognitive' cerebellum and lobules V–VI primarily represent the sensory-motor cerebellum (Kelly and Strick, 2003; Krienen and Buckner, 2009; O'Reilly et al., 2010). The between-group difference in regional homogeneity was located in motor territories of the cerebellar hemisphere, namely the cerebellar lobules V and VI that are connected with cortical motor regions. Motor impairment is prevalent in MS and is mainly caused by lesions located along the CST, in particular in the spinal cord. However, we found no relation between regional homogeneity of the cerebellum and CST lesion load. This negative finding has, however, to be interpreted with caution because our scanning protocol did not cover the spinal cord where local lesions significantly contribute to motor disability.

4.2. Local resting-state connectivity and clinical disability

Regional homogeneity of resting-state BOLD-signal fluctuations in the cerebellum correlated negatively with clinical disability as reflected by the EDSS score. Surprisingly, it was in cognitive (e.g. Crus I and Crus II) rather than motor territories of the cerebellum where the decrease in local functional connectivity was more pronounced in patients with higher disability scores. We attribute this preferential association between disease-related disability and reduced regional connectivity in cognitive areas of the cerebellar cortex to the 'resting-state' during which the fMRI data were recorded. The 'resting-state' represents a more introspective state in which cognitive operations like day-dreaming take place (Mason et al., 2007), while the motor system is idling. In fact, specifically Crus I has shown to be functionally connected with a network of cortical regions similar to the default-mode network (Krienen and Buckner, 2009) which is a network associated with intro-spection, remembering the past and planning the future (Buckner and Vincent, 2007). If this explanation were correct, an association between a reduction in regional functional connectivity in the motor cerebellum and disease-related disability might become more obvious during a task requiring sustained motor activity (e.g. tonic contraction or tapping movements).

A negative linear relationship was also found between regional homogeneity and the amount of ataxia: The stronger ataxia, the lower was local resting-state connectivity in the cerebellum. Ataxia was a frequent symptom in the patient group. 80% of the patients suffered from ataxia which is in good agreement with the relative frequency of ataxia reported in the literature (Kurtzke, 1970; Swingler and Compston, 1992). Since the EDSS also includes the assessment of cere-bellar dysfunction, it is not surprising that the ataxia and EDSS scores were highly correlated. Despite of the tight relationship between these two clinical measures, the cerebellar cluster showing a correlation between reduced local resting-state connectivity and ataxia showed little spatial overlap with the cluster showing a correlation between local resting-state connectivity and EDSS. This raises the possibility that certain 'symptom clusters' might be associated with different spatial pat-terns of disrupted local resting-state connectivity in the cerebellum. This possibility remains to be examined more systematically in future studies.

The left cerebellar cluster showing a stronger decrease in local con-nectivity with increasing disability extended into the dentate nucleus. From a circuit perspective, the dentate nucleus is located downstream to the cerebellar cortex receiving its input from the Purkinje cells. Given that the dentate nucleus is the major cerebellar output channel projecting to primary motor, premotor, oculomotor, prefrontal, and posterior parietal areas of the cortex via synapses in the thalamus and basal ganglia (Dum and Strick, 2003; Hoshi et al., 2005), the dentate nucleus is in a particularly strategic position to transmit the negative consequences of aberrant cerebellar processing to the neocortical networks. This might explain why disrupted regional integration in the dentate nucleus was significantly related to clinical disability. An ad-ditional link between an alteration of the dentate nucleus and clinical disability has recently been reported in a structural MRI study which focused on regional T2* hypointensities suggestive of pathologic iron deposition in MS (Tjoa et al., 2005). Regional T2* hypointensity in dentate nucleus was found to be the only variable that correlated with ambulatory impairment and the EDSS score.

4.3. Local versus long-range network connectivity

Outside the cerebellum, no brain region showed a significant alter-ation in regional homogeneity or a correlation between regional homoge-neity and clinical disability. By applying a whole-brain network analysis to the same rs-fMRI data set, we have previously shown that several sub-cortical clusters in the cerebral hemisphere located in the thalamus and basal ganglia have a stronger long-range connectivity with the motor resting-state network in the same group of MS patients (Dogonowski

et al., 2012). This up-regulation of motor long-range motor connectivity in the basal ganglia and thalamus as revealed by the whole-brain network analysis was not accompanied by alterations in local functional connectiv-ity. Likewise, the changes in local resting-state connectivity in the motor part of the cerebellum were not associated with alterations in long-range motor resting-state connectivity. Together, the complementary analyses of resting-state functional connectivity prompt two conclusions. From a methodological point, it can be concluded that alterations in long-range and local resting-state connectivity are not two sides of the same coin, and therefore can well dissociate. While long-range connectivity re-flects communication between distinct network nodes, local resting-state connectivity do tell something about how fine grained local neural activity is temporo-spatially correlated within a network node.

With respect to the pathophysiology of multiple sclerosis, the data indicate that cerebral subcortical structures and the cerebellum show different alterations in resting-state connectivity: While the basal gan-glia and thalamus show an expanded long-range motor connectivity, changes in local functional connectivity seem to prevail in the cerebel-lum. This study suggests a possible link between lesion load of the cer-ebellar peduncles and changes in local connectivity taking place in the cerebellum. The lesioned cerebellar peduncles in MS might become a bottleneck causing deficient parallel information transfer into the cere-bellum reflected by a reduction in cerebellar local functional connectiv-ity. The more expanded long-range connectivity of the basal ganglia and thalamus might be indicative of an impaired funnelling function of the basal ganglia and thalamus secondary to a scattered disruption of cortico-thalamic and cortico-striatal inputs due to hemispheric lesions (Dogonowski et al., 2012).

4.4. Methodological considerations

In this study, we used a relatively long rs-fMRI session collecting a comprehensive resting-state data set containing 480 volumes (20 min. acquisition) for each subject. However, we were able to reproduce the main results when only using the first half of the resting-state fMRI data (i.e., the first 238 volumes of a single session). This suggests that the sensitivity of the rs-fMRI approach might not necessarily increase when raising the number of samples in a session. Especially with respect to the long recording time, it might have been preferable to record resting-state fMRI with the subjects having the eyes open and monitoring of eye movements and electroencephalographic activity to capture any re-duction in vigilance. Another limitation of this study is that no cognitive tests were performed. This prevented us from relating local connectivity changes in the 'cognitive' cerebellum to cognitive impairments. The local connectivity changes in the right cerebellum only reached trend sig-nificance. A larger sample size might have increased the chance to yield significant changes in local connectivity also in the right cerebellum. The cross-sectional study design also has some inherent limitations. Bet-ter insights into the dynamics of functional brain connectivity changes in MS are expected from longitudinal resting-state fMRI studies which allow to trace the dynamic expression of impaired local and long-range resting-state connectivity during the course of MS and to relate these to the dynamics of clinical impairment.

We were able to relate cerebellar reductions in local connectivity with lesion load of the cerebellar peduncles. In future studies, it would be interesting to estimate the radial extension of each T2-weighted le-sion perpendicular to the peduncle axis or to test impairment of ana-tomical connectivity in the peduncles with diffusion MRI. This would give a better picture of the extent of impairment of the input and output structures of the cerebellum and might help to establish a more causal link between structural damage of the cerebellar peduncles and local changes in cerebellar resting-state connectivity.

Competing interests

Acknowledgements

References

Anderson, V.M., Fisniku, L.K., Thompson, A.J., Miller, D.H., Altmann, D., 2009. MRI measures show significant cerebellar gray matter volume loss in multiple sclerosis and are associated with cerebellar dysfunction. Mult. Scler. 15, 811–817.

Ashburner, J., Friston, K.J., 2005. Unified segmentation. Neuroimage 26, 839–851.

Birn, R.M., Diamond, J.B., Smith, M.A., Bandettini, P.A., 2006. Separating respiratory-variation-related fluctuations from neuronal-activity-related fluctuations in fMRI. Neuroimage 31, 1536–1548.

Blinkenberg, M., Rune, K., Jensen, C.V., Ravnborg, M., Kyllingsbæk, S., Holm, S., Paulson, O., Sørensen, P., 2000. Cortical cerebral metabolism correlates with MRI lesion load and cognitive dysfunction in MS. Neurology 54, 558–564.

Buckner, R.L., Vincent, J.L., 2007. Unrest at rest: default activity and spontaneous network correlations. Neuroimage 37, 1091–1096.

Derache, N., Marié, R.-M., Constans, J.-M., Defer, G.-L., 2006. Reduced thalamic and cerebellar rest metabolism in relapsing–remitting multiple sclerosis, a positron emission tomography study: correlations to lesion load. J. Neurol. Sci. 245, 103–109.

Diedrichsen, J., Balsters, J.H., Flavell, J., Cussans, E., Ramnani, N., 2009. A probabilistic MR atlas of the human cerebellum. Neuroimage 46, 39–46.

Dogonowski, A.-M., Siebner, H.R., Soelberg, P., Wu, X., Biswal, B., Paulson, O.B., Dyrby, T.B., Skimminge, A., Blinkenberg, M., Madsen, K.H., 2012. Expanded functional coupling of subcortical nuclei with the motor resting-state network in multiple sclerosis. Mult. Scler. 19, 559–566.

Dum, R.P., Strick, P.L., 2003. An unfolded map of the cerebellar dentate nucleus and its projections to the cerebral cortex. J. Neurophysiol. 89, 634–639.

Friston, K.J., Williams, S., Howard, R., Frackowiak, R.S., Turner, R., 1996. Movement-related effects in fMRI time-series. Magn. Reson. Med. 35, 346–355.

Gilmore, C.P., Donaldson, I., Bö, L., Owens, T., Lowe, J., Evangelou, N., 2009. Regional variations in the extent and pattern of grey matter demyelination in multiple sclerosis: a comparison between the cerebral cortex, cerebellar cortex, deep grey matter nuclei and the spinal cord. J. Neurol. Neurosurg. Psychiatry 80, 182–187.

Glover, G.H., Li, T.Q., Ress, D., 2000. Image-based method for retrospective correction of physiological motion effects in fMRI: RETROICOR. Magn. Reson. Med. 44, 162–167.

Hawellek, D.J., Hipp, J.F., Lewis, C.M., Corbetta, M., Engel, A.K., 2011. Increased functional connectivity indicates the severity of cognitive impairment in multiple sclerosis. Proc. Natl. Acad. Sci. U. S. A. 108, 19066–19071.

He, Y., Wang, L., Zang, Y., Tian, L., Zhang, X., Li, K., Jiang, T., 2007. Regional coherence changes in the early stages of Alzheimer's disease: a combined structural and resting-state functional MRI study. Neuroimage 35, 488–500.

Hoshi, E., Tremblay, L., Féger, J., Carras, P.L., Strick, P.L., 2005. The cerebellum communicates with the basal ganglia. Nat. Neurosci. 8, 1491–1493.

Hua, K., Zhang, J., Wakana, S., Jiang, H., Li, X., Reich, D.S., Calabresi, P.A., Pekar, J.J., van Zijl, P.C., Mori, S., 2008. Tract probability maps in stereotaxic spaces: analyses of white matter anatomy and tract-specific quantification. NeuroImage 39, 336–347.

Kelly, R.M., Strick, P.L., 2003. Cerebellar loops with motor cortex and prefrontal cortex of a nonhuman primate. J. Neurosci. 23, 8432–8444.

Krienen, F.M., Buckner, R.L., 2009. Segregated fronto-cerebellar circuits revealed by intrinsic functional connectivity. Cereb. Cortex 19, 2485–2497.

Kurtzke, J.F., 1970. Clinical manifestations of multiple sclerosis. In: Vinken, P.J., Bruyn, G.W. (Eds.), Handbook of Clinical Neurology, Amsterdam, pp. 161–216.

Kurtzke, J.F., 1983. Rating neurologic impairment in multiple sclerosis: an expanded disability status scale (EDSS). Neurology 33, 1444–1452.

Kutzelnigg, A., Faber-Rod, J.C., Bauer, J., Lucchinetti, C.F., Sorensen, P.S., Laursen, H., Stadelmann, C., Brück, W., Rauschka, H., Schmidbauer, M., Lassmann, H., 2007. Widespread demyelination in the cerebellar cortex in multiple sclerosis. Brain Pathol. 17, 38–44.

Liang, P., Liu, Y., Jia, X., Duan, Y., Yu, C., Qin, W., Dong, H., Ye, J., Li, K., 2011. Regional homogeneity changes in patients with neuromyelitis optica revealed by resting-state functional MRI. Clin. Neurophysiol. 122, 121–127.

Lowe, M.J., Phillips, M.D., Lurito, J.T., Mattson, D., Dzemidzic, M., Mathews, V.P., 2002. Multiple sclerosis: low-frequency temporal blood oxygen level-dependent fluctuations indicate reduced functional connectivity initial results. Radiology 224, 184–192.

Lund, T.E., Madsen, K.H., Sidaros, K., Luo, W., Nichols, T.E., 2006. Non-white noise in fMRI: does modelling have an impact? Neuroimage 29, 54–66.

Mason, M.F., Norton, M.I., Van Horn, J.D., Wegner, D.M., Grafton, S.T., Macrae, C.N., 2007. Wandering minds: the default network and stimulus-independent thought. Science 315, 393–395.

O'Reilly, J.X., Beckmann, C.F., Tomassini, V., Ramnani, N., Johansen-Berg, H., 2010. Distinct and overlapping functional zones in the cerebellum defined by resting state functional connectivity. Cereb. Cortex 20, 953–965.

Oldfield, R.C., 1971. Assessment and analysis of handedness—Edinburgh inventory. Neuropsychologia 9, 97–113.

Polman, C.H., Reingold, S.C., Edan, G., Filippi, M., Hartung, H.P., Kappos, L., Lublin, F.D., Metz, L.M., McFarland, H.F., O'Connor, P.W., Sandberg-Wollheim, M., Thompson, A.J.,

Weinshenker, B.G., Wolinsky, J.S., 2005. Diagnostic criteria for multiple sclerosis: 2005 revisions to the "McDonald Criteria". Ann. Neurol. 58, 840–846.

Ramnani, N., 2006. The primate cortico-cerebellar system: anatomy and function. Nat. Rev. Neurosci. 7, 511–522.

Ravnborg, M., Grønbech-Jensen, M., Jønsson, A., 1997. The MS Impairment Scale: a pragmatic approach to the assessment of impairment in patients with multiple sclerosis. Mult. Scler. 3, 31–42.

Rocca, M.A., Absinta, M., Valsasina, P., Ciccarelli, O., Marino, S., Rovira, A., Gass, A., Wegner, C., Enzinger, C., Korteweg, T., 2009. Abnormal connectivity of the sensorimotor network in patients with MS: a multicenter fMRI study. Hum. Brain Mapp. 30, 2412–2425.

Rocca, M.A., Valsasina, P., Absinta, M., Riccitelli, G., Rodegher, M.E., Misci, P., Rossi, P., Falini, A., Comi, G., Filippi, M., 2010. Default-mode network dysfunction and cognitive impairment in progressive MS. Neurology 74, 1252–1259.

Roosendaal, S.D., Schoonheim, M.M., Hulst, H.E., Sanz-Arigita, E.J., Smith, S.M., Geurts, J.J.G., Barkhof, F., 2010. Resting state networks change in clinically isolated syndrome. Brain 133, 1612–1621.

Saini, S., DeStefano, N., Smith, S., Guidi, L., Amato, M.P., Federico, A., Matthews, P.M., 2004. Altered cerebellar functional connectivity mediates potential adaptive plasticity in patients with multiple sclerosis. J. Neurol. Neurosurg. Psychiatry 75, 840–846.

Swingler, R.J., Compston, D.A., 1992. The morbidity of multiple sclerosis. Q. J. Med. 83, 325–337.

Tjoa, C.W., Benedict, R.H.B., Weinstock-Guttman, B., Fabiano, A.J., Bakshi, R., 2005. MRI T2 hypointensity of the dentate nucleus is related to ambulatory impairment in multiple sclerosis. J. Neurol. Sci. 234, 17–24.

Trapp, B.D., Peterson, J., Ransohoff, R.M., Rudick, R., Mörk, S., Bö, L., 1998. Axonal transection in the lesions of multiple sclerosis. N. Engl. J. Med. 338, 278–285.

Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., Mazoyer, B., Joliot, M., 2002. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. Neuroimage 15, 273–289.

Wu, T., Long, X., Zang, Y., Wang, L., Hallett, M., Li, K., Chan, P., 2009. Regional homogeneity changes in patients with Parkinson's disease. Hum. Brain Mapp. 30, 1502–1510.

Zang, Y., Jiang, T., Lu, Y., He, Y., Tian, L., 2004. Regional homogeneity approach to fMRI data analysis. Neuroimage 22, 394–400.

Zuo, X.N., Di Martino, A., Kelly, C., Shehzad, Z.E., Gee, D.G., Klein, D.F., Castellanos, F.X., Biswal, B.B., Milham, M.P., 2010. The oscillating brain: complex and reliable. Neuroimage 49, 1432–1445.

Identifying Modular Relations in Complex Brain Networks

Andersen, K. W., Mørup, M., Siebner, H., Madsen, K. H., and Hansen, L. K. (2012), 'Identifying modular relations in complex brain networks', in '2012 IEEE International Workshop on Machine Learning for Signal Processing', IEEE.

IDENTIFYING MODULAR RELATIONS IN COMPLEX BRAIN NETWORKS

Kasper Winther Andersen[a,b]*, Morten Mørup*[a]*, Hartwig Siebner*[b]*, Kristoffer H Madsen*[b]*, Lars Kai Hansen*[a]

[a]DTU Informatics, Technical University of Denmark.
[b]Danish Research Centre for Magnetic Resonance, Copenhagen University Hospital Hvidovre

ABSTRACT

We evaluate the infinite relational model (IRM) against two simpler alternative nonparametric Bayesian models for identifying structures in multi subject brain networks. The models are evaluated for their ability to predict new data and infer reproducible structures. Prediction and reproducibility are measured within the data driven NPAIRS split-half framework. Using synthetic data drawn from each of the generative models we show that the IRM model outperforms the two competing models when data contain relational structure. For data drawn from the other two simpler models the IRM does not overfit and obtains comparable reproducibility and predictability. For resting state functional magnetic resonance imaging data from 30 healthy controls the IRM model is also superior to the two simpler alternatives, suggesting that brain networks indeed exhibit universal complex relational structure in the population.

Index Terms— Infinite Relational Model, Complex Networks, fMRI

1. INTRODUCTION

The brain is composed of about 10^{11} neurons connected by more than 10^{14} synapses forming an almost unfathomable complex network [1]. Neuroimaging, e.g., by functional magnetic resonance imaging (fMRI), has the potential to identify structures in the brain network. fMRI allows one to measure the blood oxygen level and thereby an indirect measure of the neuronal activity in a relatively fine spatial resolution. As the signal-to-noise ratio in fMRI is rather poor we are in many cases forced to average within a population of subjects to obtain reproducible results. This averaging further limits the spatial resolution of our models as considerable individual variability exist in both anatomy and function [2].

While the dominant paradigm in functional neuroimage analysis has been to identify local components of the network based on the functional segregation hypothesis, see e.g. [3], there is a current move towards modeling more global properties of the brain network including both functional and structural aspects, for reviews see, e.g., [4, 5]. Functional brain network models typically summarize temporal relations be-

tween regions of the brain whereas structural networks are obtained by, e.g., using tractography methods to track the white matter bundles connecting different brain regions. One of the challenges in brain imaging is to understand how brain modules identified under the functional segregation approach work together to perform important information processing tasks such as perception, cognition, and decision making. We currently pursue a strategy which is designed to mediate between functional localization and global description [6]. The main idea is to search for so-called 'communities' of nodes characterized as groups of nodes that share certain connectivity properties, see e.g. [7] for background and references on community detection.

In neuroimaging models we typically face two equally important objectives, namely on one hand to identify predictive models, i.e., models that generalize to new data, and on the other hand that the structures we learn are trustworthy, e.g., being reproducible across different subject groups. The two objectives are quantified in the so-called NPAIRS split-half resampling framework [8]. This framework has been used for evaluation of Bayesian models before, e.g., in [9] in which generalizability and reproducibility of the parameter posterior distribution were of interest, however, the framework has not been used before for evaluation of community detection models. Here we for the first time investigate the predictability and reproducibility of three Bayesian nonparametric models which are able to infer community structure in complex networks.

The models are the infinite relational model (IRM) [10, 11], the infinite diagonal model (IDM), and the model proposed by [12] which will be termed infinite Hofman-Wiggins (IHW) model, the term infinite referring to the fact that the complexity in terms of the number of communities is formally unlimited. The three models vary in expressiveness, the IRM allows different link characteristics within and between components and are thereby able to infer general relations between communities, the IDM is parametrized with different link probabilities within communities but have a single common between community link probability, while the link probabilities in IHW are described with only two parameters; one within and one between component link probability.

As mentioned we are interested in detecting community

structure shared by a population of fMRI subjects. Thus while we estimate the functional networks in each individual independently, all three models represent all subjects by a common community structure and common link probabilities. An additional benefit of this approach, compared to individual models investigated [6] is that it become possible to compute test likelihood for unseen data (subjects). Also, having common link probabilities the latent variables of the model do not scale with the number of subjects, hence, allow for analysis of a large number of subjects.

In conclusion: The main contributions of this paper are i) a demonstration that a data driven predictability and reproducibility framework for evaluation of neuroimaging models can be adapted to Bayesian community detection and ii) use of the framework to test whether the expressiveness of the full IRM is needed to describe the functional networks of the resting state as measured by fMRI.

2. METHODS

Data in this paper consists of N undirected graphs, one per subject, defined on a common set of nodes and described by individual adjacency matrices $\mathbf{A}^{(n)}$ for subject n. $A_{i,j}^{(n)} = 1$ if nodes i and j have similar hemodynamic responses and zero otherwise.

2.1. Models

Formally we write the three generative models as:

$$
\begin{aligned}
\mathbf{Z}|\alpha &\sim \mathrm{CRP}(\alpha) \\
\rho_{a,b}|\boldsymbol{\beta}^+, \boldsymbol{\beta}^- &\sim \mathrm{Beta}(\beta_{a,b}^+, \beta_{a,b}^-) \\
A_{i,j}^{(n)}|\mathbf{Z}, \boldsymbol{\rho} &\sim \mathrm{Bernoulli}(\mathbf{z}_{i_r}\boldsymbol{\rho}\mathbf{z}_{j_r}^\top).
\end{aligned}
$$

$\mathbf{Z}$ is the $[J \times D]$ binary matrix indicating group membership for each node and have a prior distribution given by the Chinese Restaurant Process (CRP) with the hyperprior parameter α. $\mathbf{z}_{i_r}$ is the ith row vector of $\mathbf{Z}$. $\boldsymbol{\rho}$ is a symmetric matrix indicating the probability of links between each pair of components. Here $\boldsymbol{\rho}$ is shared across graphs. The $\boldsymbol{\rho}$ structure define the difference between the models compared in this paper, which will be described below. The link probabilities have a prior given by the Beta function with the two hyperprior parameters $\beta_{a,b}^+$ and $\beta_{a,b}^-$ giving the pseudocount of links and nonlinks between components a and b respectively. The adjacency matrix $\mathbf{A}^{(n)}$ for graph n has the prior given by a Bernoulli trial biased by the link probability of links between the components in which the two nodes participate. Three different structures for the link probability matrix are considered. The full IRM model having a full upper triangular matrix, the infinite diagonal model (IDM) having unique elements in the diagonal and identical off-diagonal elements, and the model proposed in [12] having two parameters describing

the within and between component link probabilities, respectively.

$$
\boldsymbol{\rho}_{\mathrm{IRM}} = \begin{bmatrix} \rho_{11} & \rho_{12} & \cdots & \rho_{1m} \\ \rho_{21} & \rho_{22} & \cdots & \rho_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{m1} & \rho_{m2} & \cdots & \rho_{mm} \end{bmatrix} \quad (1)
$$

$$
\boldsymbol{\rho}_{\mathrm{IDM}} = \begin{bmatrix} \rho_1 & \rho_0 & \cdots & \rho_0 \\ \rho_0 & \rho_2 & \cdots & \rho_0 \\ \vdots & \vdots & \ddots & \vdots \\ \rho_0 & \rho_0 & \cdots & \rho_m \end{bmatrix} \quad (2)
$$

$$
\boldsymbol{\rho}_{\mathrm{IHW}} = \begin{bmatrix} \rho_c & \rho_0 & \cdots & \rho_0 \\ \rho_0 & \rho_c & \cdots & \rho_0 \\ \vdots & \vdots & \ddots & \vdots \\ \rho_0 & \rho_0 & \cdots & \rho_c \end{bmatrix} \quad (3)
$$

The joint likelihood over graphs can be written as

$$
\begin{aligned}
&P(\mathbf{A}^{(1)}, ..., \mathbf{A}^{(N)}|\mathbf{Z}, \boldsymbol{\rho}) \\
&= \prod_n \prod_{j>i} (\mathbf{z}_{i_r}\boldsymbol{\rho}\mathbf{z}_{j_r}^\top)^{A_{i,j}^{(n)}} (1 - \mathbf{z}_{i_r}\boldsymbol{\rho}\mathbf{z}_{j_r}^\top)^{\left(1 - A_{i,j}^{(n)}\right)} \\
&= \prod_{j \geq i} (\mathbf{z}_{i_r}\boldsymbol{\rho}\mathbf{z}_{j_r}^\top)^{\left(\sum_n A_{i,j}^{(n)}\right)} (1 - \mathbf{z}_{i_r}\boldsymbol{\rho}\mathbf{z}_{j_r}^\top)^{\left(N - \sum_n A_{i,j}^{(n)}\right)} \quad (4)
\end{aligned}
$$

We note that the joint likelihood can be written efficiently using the aggregate adjacency matrix, $\mathbf{A}^{\mathrm{tot}} = \sum_n \mathbf{A}^{(n)}$.

2.1.1. IRM Model Inference

In this section we derive the posterior likelihood of a node's assignment given the assignments of the remaining nodes, for the IRM model, which is needed in the model inference. For the two other models please refer to [7]. By integrating out $\boldsymbol{\rho}$ from eq (4) the posterior can be written as

$$
\begin{aligned}
&P(\mathbf{A}^{\mathrm{tot}}|\mathbf{Z}, \boldsymbol{\beta}^+, \boldsymbol{\beta}^-) \\
&= \int P(\mathbf{A}^{\mathrm{tot}}|\mathbf{Z}, \boldsymbol{\rho})P(\boldsymbol{\rho}|\boldsymbol{\beta}^+, \boldsymbol{\beta}^-)\mathrm{d}\boldsymbol{\rho} \\
&= \prod_{a \geq b} \frac{\mathrm{Beta}(M_{a,b}^+ + \beta_{a,b}^+, M_{a,b}^- + \beta_{a,b}^-)}{\mathrm{Beta}(\beta_{a,b}^+, \beta_{a,b}^-)},
\end{aligned}
$$

where $M_{a,b}^+$ is the total number of links and $M_{a,b}^+$ is the total number of non-links between the groups a and b across all graphs. Using Bayes' theorem the posterior likelihood can be found as

$$
\begin{aligned}
&P(\mathbf{Z}|\mathbf{A}^{(1)}, ..., \mathbf{A}^{(N)}, \boldsymbol{\beta}^+, \boldsymbol{\beta}^-, \alpha) \propto \\
&P(\mathbf{A}^{\mathrm{tot}}|\mathbf{Z}, \boldsymbol{\beta}^+, \boldsymbol{\beta}^-)P(\mathbf{Z}|\alpha) = \\
&\prod_{a \geq b} \frac{\mathrm{Beta}(M_{a,b}^+ + \beta_{a,b}^+, M_{a,b}^- + \beta_{a,b}^-)}{\mathrm{Beta}(\beta_{a,b}^+, \beta_{a,b}^-)} \\
&\times \left[\alpha^D \frac{\Gamma(\alpha)}{\Gamma(J + \alpha)} \prod_a \Gamma(n_a)\right],
\end{aligned}
$$

where n_a is the number of nodes assigned to component a. The posterior likelihood for a node's assignment given the assignment of the remaining nodes is given by

$$P(Z_{i,a} = 1 | \mathbf{Z} \backslash \mathbf{z}_{i_r}, \mathbf{A}^{(1)}, ..., \mathbf{A}^{(N)}) \propto$$
$$\begin{cases} m_a \prod_b \frac{\text{Beta}(M_{a,b}^+ + \beta_{a,b}^+, M_{a,b}^- + \beta_{a,b}^-)}{\text{Beta}(\beta_{a,b}^+, \beta_{a,b}^-)} & \text{if } m_a > 0 \\ \alpha \prod_b \frac{\text{Beta}(M_{a,b}^+ + \beta_{a,b}^+, M_{a,b}^- + \beta_{a,b}^-)}{\text{Beta}(\beta_{a,b}^+, \beta_{a,b}^-)} & \text{otherwise.} \end{cases} \quad (5)$$

$m_a = \sum_{j \neq i} \mathbf{Z}(j, a)$ is the size of the ath functional group disregarding the assignment of the ith node. This posterior likelihood can be evaluated efficiently since we only need $\mathbf{M}^+$ and $\mathbf{M}^-$ and further to evaluate the Beta function for entries affected by the considered assignment change.

2.2. NPAIRS Evaluation Criteria

The models' predictability is evaluated using test log likelihood. The data is randomly split in 2 equally sized sets (S1 and S2) and inference is made separately on each set. The node assignment and link probabilities from the MAP solution are used to calculate the test log likelihood of the other (unseen) split. The average test log likelihood of the two splits is used as the predictability measure. The log likelihood of set S2 for the model inferred using S1 is given as

$$\log L(\mathbf{Z}, \boldsymbol{\rho} | \mathbf{A}^{S2,(1)}, ..., \mathbf{A}^{S2,(N)}) =$$
$$\frac{1}{N} \sum_{n=1}^{N} \sum_{j > i} \left[A_{i,j}^{S2,(n)} \log(\mathbf{z}_{i_r} \boldsymbol{\rho} \mathbf{z}_{j_r}^\top) + (1 - A_{i,j}^{S2,(n)}) \log(1 - \mathbf{z}_{i_r} \boldsymbol{\rho} \mathbf{z}_{j_r}^\top) \right]$$

Likewise, the reproducibility of the identified community structures is measured by the mutual information (MI) between the node assignments $\mathbf{Z}^{(S1)}$ and $\mathbf{Z}^{(S2)}$ of the MAP solution for each split,

$$\text{MI}(\mathbf{Z}^{(S1)}, \mathbf{Z}^{(S2)}) =$$
$$\sum_{i=1}^{D1} \sum_{j=1}^{D2} p(\mathbf{z}_{i_r}^{(S1)}, \mathbf{z}_{j_r}^{(S2)}) \log \left(\frac{p(\mathbf{z}_{i_r}^{(S1)}, \mathbf{z}_{j_r}^{(S2)})}{p(\mathbf{z}_{i_r}^{(S1)}) p(\mathbf{z}_{j_r}^{(S2)})} \right).$$

3. EXPERIMENTS AND RESULTS

In this section we present results for two different dataset. The first data set is synthetic data generated from each of the three different models. The second dataset is a dataset of resting state fMRI data from 30 healthy controls.

Model inference is based on a Gibbs sampling scheme in combination with split-merge Metropolis-Hastings updates [6, 10, 13]. Each node is initially assigned to one of 50 groups at random. The algorithm runs for 500 iterations, where at each iteration a Gibbs sampling scan is followed by a Split-Merge step. In the model inference $\alpha = \log(J)$ and $\boldsymbol{\beta}^+ = \boldsymbol{\beta}^- = 1$.

3.1. Synthetic Data

Data from each of the three different generative models, IRM, IDM, and IHW, were generated. Assignments for $J = 100$ nodes was drawn from the CRP with $\alpha = 5$. Then the link probabilities between components was drawn from the Beta function using

$$\beta_{a,b}^+ = \begin{cases} 2 \text{ if a=b} \\ 1 \text{ otherwise} \end{cases} \quad \text{and} \quad \beta_{a,b}^- = \begin{cases} 3 \text{ if a=b} \\ 5 \text{ otherwise.} \end{cases}$$

constrained according to equations (1-3) for the three models. Then, using the component assignments and link probabilities 20 different adjacency matrices (graphs) per dataset was drawn using the Bernoulli function. For testing the models' predictability and reproducibility 500 different half-splits per dataset was generated. For each split each of the models was inferred and the predictability and reproducibility was evaluated as described in section 2.2.

Reproducibility-predictability plots are shown in Fig. 1 and the histogram of the number of estimated components are shown in Fig. 2. From Fig. 1 it is evident that for the IRM data with complex relations the IRM is superior in estimating the complex component relations as seen by having better reproducibility and predictability. The two other models produce comparable results with the IDM model having slightly better reproducibility. The IRM model is close in estimating the true number of components while the two other models underestimates the number of component with a factor 2.

For the IDM generated data the IDM model has better predictability and reproducibility than the IRM model, while the IHW model obtains slightly better reproducibility than the IDM model but with worse predictability. The better reproducibility of the IHW model should be seen in the light of overestimated number of components which increases the mutual information if the assignments of the between splits agree. The IRM underestimate the true number of components, which also cause the lower reproducibility.

For the IHW generated data all three models have comparable results with the IHW model having slightly better reproducibility than the IDM which is turn is slightly better than the IRM model. The IDM and IHW estimates the true number of components well, while the IRM model underestimates the true number of components.

3.2. Resting state fMRI data

Resting state functional magnetic resonance imaging (rs-fMRI) data from $N = 30$ healthy controls was recorded for 20 min (482 volumes) per subject. The first two volumes were discarded to account for T1 equilibrium effects, the remaining 480 volumes were realigned to the time-series mean
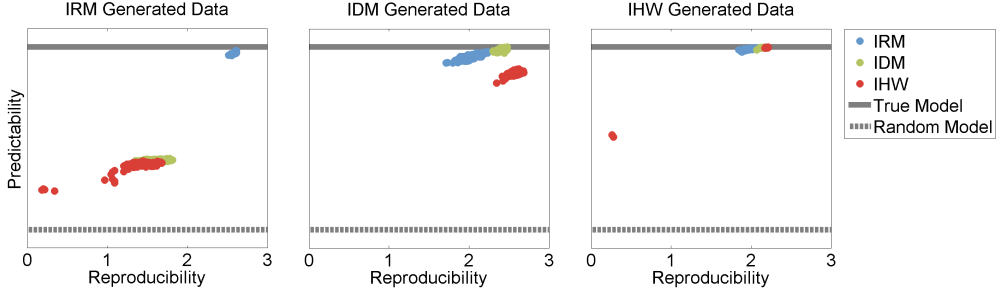
Fig. 1. Predictability (test log likelihood) as function of reproducibility (mutual information) for data generated with the IRM, IDM, and IHW models respectively. Solid lines indicate the log likelihood for the true model (which generated the data). Dotted lines indicate log likelihood for a random model in which all elements in the ρ matrix are identical. For the IRM generated data the IRM model is superior compared to the two other models in both the predictability and reproducibility measures with the two other models having comparable results. For the IDM generated data the IDM is slightly better better than the IRM model and has a better predictability than the IHW, though with the IHW reproducing slightly better. For the IHW generated data all models are almost comparable though with slightly better results in favor of the simpler models.
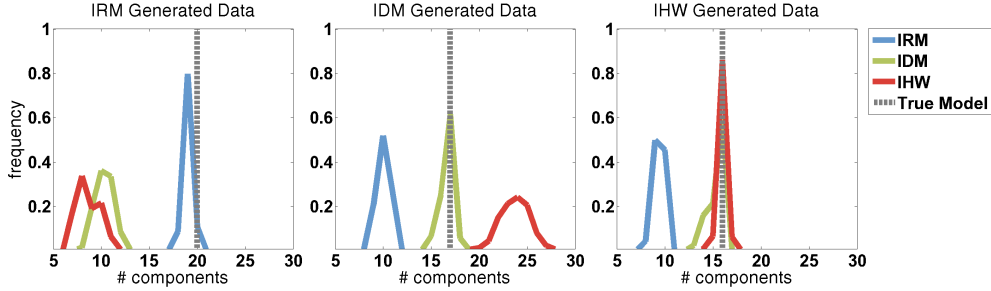
Fig. 2. Histogram of number of components estimated using the three models. Data was generated using either the IRM, IDM, or IHW generative model. For the IRM generated data the IRM model is close in estimating the true number of components (indicated with the vertical dotted line) while the two other models underestimate the number of components. For the IDM generated data the IDM model closely finds the right number of components while the IRM model underestimate and IHW overestimate the number of components. For the IHW generated data both the IDM and IHW finds the true number of components while the IRM underestimate.

and spatially normalized to the MNI template using SPM. Nuisance effects related to residual movement or physiological effects were removed using a linear filter comprised of 24 motion related and a total of 64 physiological effects including cardiac, respiratory, respiration volume over time, and time series from left and right hemispheres CSF and white matter voxels.

The mean signal in each of the 116 regions covered in the AAL database [14] was extracted and the $[116 \times 116]$ correlation matrix was created for each subject. These matrices were thresholded to include the top 1000 links and thereby formed the adjacency matrices. 200 different half-splits were gener-

ated and the reproducibility and predictability were calculated between the half-splits for each of the three models.

For visualization of the estimated communities a co-occurrence matrix of the count of each pair of nodes were in the same community were made. Based on this co-occurrence matrix we performed agglomerative hierarchical clustering using average linkage. This clustering was thresholded to include the number of communities corresponding to the median number found across the 400 half-splits. Fig. 3 shows the reproducibility-predictability plot and Fig. 4 plots the number of components estimated using each of the three models. The more complex IRM model interestingly yields better repro-
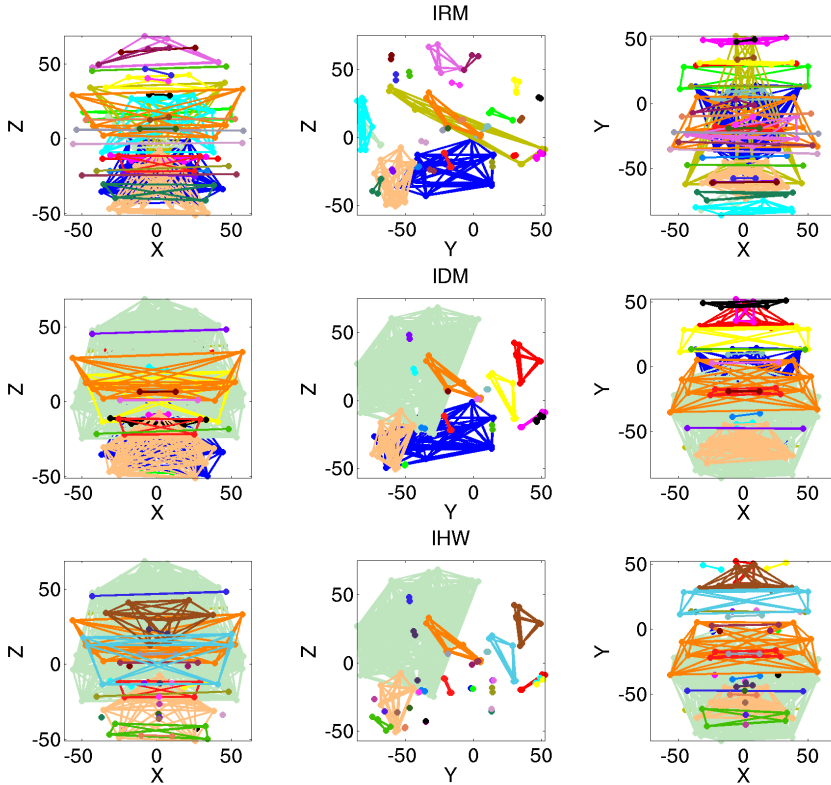
Fig. 5. Resting state fMRI data. Community structure for each of the three models. Top row IRM model, middle row IDM model, bottom row IHW model. For each model the consensus of 400 runs are plotted at three different planes in the 3 dimensional stereotaxic MNI space (X: left to right; Y: posterior to anterior; Z: ventral to dorsal).

ducibility and predictability compared with the other models. The IDM and IHW models yields comparable predictability but the IHW shows better reproducibility. Fig. 5 shows the layout of the identified communities in the stereotaxic MNI space in coronal, sagittal, and axial planes respectively. Lines connecting nodes in a community are shown with the same color. For all three models the communities are highly symmetric across hemispheres. The two simpler models IDM and IHW produce a huge cluster in the occipital and parietal lobes, while especially the IHW produces a large number of small communities. The huge component is segregated in several components with the IRM model, made possible by the more expressive parametrization.

4. DISCUSSION AND CONCLUSION

We used the NPAIRS framework for quantification of the predictability and reproducibility of three models for structural inference in complex brain networks. For synthetic data generated from each of the models we show that the IRM model identifies a general relational structure when such exist in network data. The two simpler models fail in both predicting unseen data and reproduce across datasets. When data is generated using the simpler models the IRM still produce comparable predictability and reproducibility results and thus does not seem to produce severe overfits.

Based on resting state fMRI data from 30 subjects extracted in 116 brain regions covering the whole brain we show that the IRM model is superior in terms of both predictability and reproducibility compared with the less expressive models. Visual inspection of the inferred communities indeed indicates that the richer representation of the IRM produces a more detailed yet reproducible connectivity pattern in the resting human brain.

Fig. 3. Reproducibility-predictability plot of the fMRI dataset. Dotted line indicates log likelihood for a random model in which all elements in the ρ matrix are identical.

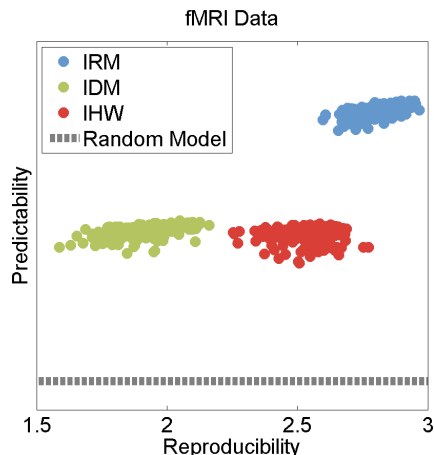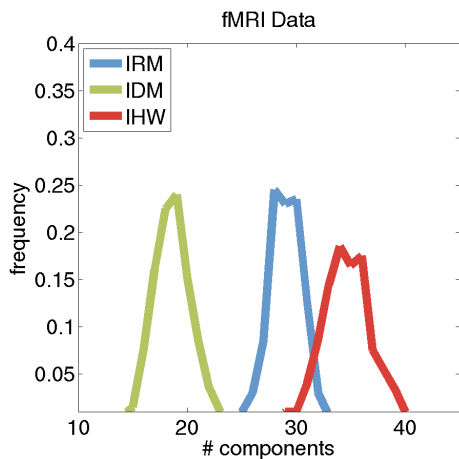**Fig. 4**. Number of components estimated using the IRM, IDM, and IHW models for the 400 half-splits of the data.

Acknowledgement

5. REFERENCES

[1] S.M. Blinkov and I.I. Glezer, *The human brain in figures and tables: a quantitative handbook*, Basic Books New York:, 1968.

[2] A.M. Galaburda, G.D. Rosen, and G.F. Sherman, "Individual variability in cortical organization: its relationship to brain laterality and implications to function," *Neuropsychologia*, vol. 28, no. 6, pp. 529–546, 1990.

[3] K.J. Friston, A.P. Holmes, K.J. Worsley, J.P. Poline, C.D. Frith, and R.S.J. Frackowiak, "Statistical parametric maps in functional imaging: a general linear approach," *Human brain mapping*, vol. 2, no. 4, pp. 189–210, 1994.

[4] Olaf Sporns, "The human connectome: a complex network.," *Annals of the New York Academy of Sciences*, vol. 1224, no. 1, pp. 109–25, Apr. 2011.

[5] Edward T Bullmore and Danielle S Bassett, "Brain graphs: graphical models of the human brain connectome.," *Annual review of clinical psychology*, vol. 7, pp. 113–40, Apr. 2011.

[6] M Mørup, K.H. Madsen, A.M. Dogonowski, H. Siebner, and L.K. Hansen, "Infinite Relational Modeling of Functional Connectivity in Resting State fMRI," *Neural Information Processing Systems 23*, 2010.

[7] M. Mørup and M. N. Schmidt, "Bayesian community detection," *Neural Computation*, 2012.

[8] S.C. Strother, J. Anderson, L.K. Hansen, U. Kjems, R. Kustra, J. Sidtis, S. Frutiger, S. Muley, S. LaConte, and D. Rottenberg, "The quantitative evaluation of functional neuroimaging experiments: The npairs data analysis framework," *NeuroImage*, vol. 15, no. 4, pp. 747–771, 2002.

[9] D.J. Jacobsen, L.K. Hansen, and K.H. Madsen, "Bayesian model comparison in nonlinear bold fmri hemodynamics," *Neural computation*, vol. 20, no. 3, pp. 738–755, 2008.

[10] Charles Kemp, J.B. Tenenbaum, T.L. Griffiths, T. Yamada, and N. Ueda, "Learning systems of concepts with an infinite relational model," in *Proceedings of the National Conference on Artificial Intelligence*. 2006, vol. 21, pp. 381–388, Menlo Park, CA; Cambridge, MA; London; AAAI Press; MIT Press; 1999.

[11] Zhao Xu, Volker Tresp, Kai Yu, and H.P. Kriegel, "Infinite hidden relational models," *In Proceedings of the 22nd International Conference on Uncertainty in Artificial Intelligence*, 2006.

[12] Jake M Hofman and Chris H Wiggins, "A Bayesian approach to network modularity, 2008," *Phys. Rev. Lett*, vol. 100, no. 25, pp. 1–4, 2008.

[13] Sonia Jain and Radford M Neal, "A Split-Merge Markov chain Monte Carlo Procedure for the Dirichlet Process Mixture Model," *Journal of Computational and Graphical Statistics*, vol. 13, no. 1, pp. 158–182, 2004.

[14] N Tzourio-Mazoyer, B Landeau, and D Papathanassiou, "Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain," *Neuroimage*, vol. 15, pp. 273–289, 2002.

Identification of Functional Clusters in the Striatum Using Infinite Relational Modeling

Andersen, K. W., Madsen, K. H., Siebner, H., Hansen, L. K., and Mørup, M. (2012), 'Identification of Functional Clusters in the Striatum Using Infinite Relational Modeling', in G. Langs, I. Rish, M. Grosse-Wentrup and B. Murphy, eds, 'Machine Learning and Interpretation in Neuroimaging', Lecture Notes in Computer Science, Springer Berlin Heidelberg, pp. 226-233.

Identification of Functional Clusters in the Striatum Using Infinite Relational Modeling

Kasper Winther Andersen[1,2], Kristoffer Hougaard Madsen[1,2],
Hartwig Siebner[2], Lars Kai Hansen[1], and Morten Mørup[1]

[1] DTU Informatics, Technical University of Denmark
[2] Danish Research Centre for Magnetic Resonance, Copenhagen University Hospital Hvidovre
{kwjo,lkh,mm}@imm.dtu.dk,
{stoffer,hartwig.siebner}@drcmr.dk

Abstract. In this paper we investigate how the Infinite Relational Model can be used to infer functional groupings of the human striatum using resting state fMRI data from 30 healthy subjects. The Infinite Relational Model is a non-parametric Bayesian method for infering community structure in complex networks. We visualize the solution found by performing evidence accumulation clustering on the maximum a posterior solutions found in 100 runs of the sampling scheme. The striatal groupings found are symmetric between hemispheres indicating that the model is able to group voxels across hemispheres, which are involved in the same neural computations. The reproducibility of the groupings found are assessed by calculating mutual information between half splits of the subject sample for various hyperparameter values. Finally, the model's ability to predict unobserved links is assessed by randomly treating links and non-links in the graphs as missing. We find that the model is performing well above chance for all subjects.

Keywords: complex network, graph theory, infinite relational model, basal ganglia, striatum.

1   Introduction

Recently, graph theoretical network modeling has gained a lot of attention in neuroimaging, for reviews see, e.g, [3,15]. Both functional networks (using modalities such as fMRI, EEG, and MEG) and anatomical brain networks (using DWI) have been analyzed using complex network methods. These studies cover both studies of the healthy brain as well as a wide range of neuropsychiatric and neurologic disorders [16]. In this work we use the Infinite Relational Model (IRM) [8,17] to infer functional groupings of the human striatum. The IRM model is a non-parametric Bayesian network model, which assigns nodes into non-overlapping groups. The probability of a link between two nodes is determined by the groups the nodes are assigned to. During inference the number of groups and the group

assignments are inferred, while the group link probabilities can be integrated out of the model and are therefore not determined during inference. These probabilities are easily calculated afterwards given the group assignments. The IRM allows analysis of multi-graph networks and thus provides a natural framework for analyzing multiple subjects at once as demonstrated in [13].

The basal ganglia (BG) process information from the cerebral cortex in segregated parallel cortico-BG-thalamocortical loops [1]. The BG are involved in the adaptation of complex goal related behaviors [4,6] and play a key role in the pathophysiology of many neurological (e.g., Parkinsons disease) and psychiatric (e.g., schizophrenia) disorders [11,14]. The caudate nucleus and putamen (i.e., dorsal striatum) are the main input structures of the BG receiving topographically organized inputs from the cortex. Striatal sub-territory receives specific cortical inputs via corticostriatal feed-forward projections originating from largely segregated cortical input zones [6,12]. The BG anatomy and function is largely symmetric between the two hemispheres.

2   Methods

2.1   Data

Resting state functional magnetic resonance imaging (rs-fMRI) data from $N = 30$ healthy controls was recorded for 20 min (482 volumes) per subject. The first two volumes were discarded to account for T1 equilibrium effects, the remaining 480 volumes were realigned to the time-series mean and spatially normalized to the MNI template using SPM. Nuisance effects related to residual movement or physiological effects were removed using a linear filter comprised of 24 motion related and a total of 64 physiological effects including cardiac, respiratory, respiration volume over time, and time series from left and right hemispheres CSF and white matter voxels. An anatomical mask consisting of the caudate nucleus and putamen, which was made in WFU PickAtlas [10] using the Talairach Daemon atlas [9], was used to extract the time series of the $J = 825$ voxels from all subjects.

The network graph representing functional connectivity in subject n is represented by the $[J \times J]$ adjacency matrix $\mathbf{A}^{(n)}$. Each graph is then composed of J nodes and $\mathbf{A}^{(n)}(i,j)$ is 1 if a link is present between voxels i and j and 0 elsewhere. $\mathbf{A}^{(n)}$ was obtained for each subject by first calculating the upper triangular part of the Pearson correlation matrix and then thresholding the matrix to include the highest 5000 positive correlations.

2.2   The Infinite Relational Model

Following the notation in [13] the IRM generative model can be written as:

$$\mathbf{Z}|\alpha \sim \mathrm{CRP}(\alpha)$$
$$\boldsymbol{\rho}^{(n)}|\beta \sim \mathrm{Beta}(\beta, \beta)$$
$$\mathbf{A}^{(n)}(i,j)|\mathbf{Z}, \boldsymbol{\rho}^{(n)} \sim \mathrm{Bernoulli}(\mathbf{z}_{i_r} \boldsymbol{\rho}^{(n)} \mathbf{z}_{j_r}^\top).$$

As such, the probability of a link between two voxels is determined by the groups in which the voxels are members of. $\boldsymbol{\rho}^{(n)}$ is the subject specific group link probability matrix and defines the probability of links between groups. $\mathbf{Z}$ is a $[J \times D]$ binary matrix indicating group membership for each voxel and is shared across all subjects. We use symmetric Beta functions with hyperparameter β as priors for the group link probabilities and the Chinese Restaurant Process (CRP) is used as prior for the voxel group assignments. By integrating $\boldsymbol{\rho}$ out the posterior can be written as:

$$P(\mathbf{A}^{(n)}|\mathbf{Z}, \beta) = \int P(\mathbf{A}^{(n)}|\mathbf{Z}, \boldsymbol{\rho}^{(n)}) P(\boldsymbol{\rho}^{(n)}|\beta) \mathrm{d}\boldsymbol{\rho}^{(n)}$$

$$= \prod_{a \geq b} \frac{\mathrm{Beta}(\mathbf{M}_+^{(n)}(a, b) + \beta, \mathbf{M}_-^{(n)}(a, b) + \beta)}{\mathrm{Beta}(\beta, \beta)},$$

where $\mathbf{M}_+^{(n)}(a, b) = (1 - \frac{1}{2}\delta_{a,b})\mathbf{z}_a^\top(\mathbf{A}^{(n)} + \mathbf{A}^{(n)\top})\mathbf{z}_b$ is the number of links and $\mathbf{M}_-^{(n)}(a, b) = (1 - \frac{1}{2}\delta_{a,b})\mathbf{z}_a^\top(\boldsymbol{e}\boldsymbol{e}^\top - \boldsymbol{I})\mathbf{z}_b - \mathbf{M}_+^{(n)}(a, b)$ is the number of non-links between group a and b. $\boldsymbol{e}$ is a vector of length J with ones in all entries. The subjects' adjacency matrices are assumed independent thus their joint distribution is:

$$P(\mathbf{A}^{(1)}, ..., \mathbf{A}^{(N)}|\mathbf{Z}, \beta) = \prod_n \prod_{a \geq b} \frac{\mathrm{Beta}(\mathbf{M}_+^{(n)}(a, b) + \beta, \mathbf{M}_-^{(n)}(a, b) + \beta)}{\mathrm{Beta}(\beta, \beta)}.$$

Using Bayes' theorem the posterior likelihood can be found as:

$$P(\mathbf{Z}|\mathbf{A}^{(1)}, ..., \mathbf{A}^{(N)}, \beta, \alpha) \propto P(\mathbf{A}^{(1)}, ..., \mathbf{A}^{(N)}|\mathbf{Z}, \beta,) P(\mathbf{Z}|\alpha) =$$

$$\left[\prod_n \prod_{a \geq b} \frac{\mathrm{Beta}(\mathbf{M}_+^{(n)}(a, b) + \beta, \mathbf{M}_-^{(n)}(a, b) + \beta)}{\mathrm{Beta}(\beta, \beta)} \right] \left[\alpha^D \frac{\Gamma(\alpha)}{\Gamma(J + \alpha)} \prod_a \Gamma(n_a) \right].$$

For model inference we use a Gibbs sampling scheme in combination with split-merge sampling [7,8,13], requiring the posterior likelihood for a node's assignment given the assignment of the remaining nodes:

$$P(\mathbf{Z}(i, a) = 1|\mathbf{Z}\backslash\mathbf{z}_{i_r}, \mathbf{A}^{(1)}, ..., \mathbf{A}^{(N)})$$

$$\propto \begin{cases} m_a \prod_n \prod_b \frac{\mathrm{Beta}(\mathbf{M}_+^{(n)}(a,b) + \beta, \mathbf{M}_-^{(n)}(a,b) + \beta)}{\mathrm{Beta}(\mathbf{M}_{+\backslash i}^{(n)}(a,b) + \beta, \mathbf{M}_{-\backslash i}^{(n)}(a,b) + \beta))} & \text{if } m_a > 0 \\ \alpha \prod_n \prod_b \frac{\mathrm{Beta}(\mathbf{M}_+^{(n)}(a,b) + \beta, \mathbf{M}_-^{(n)}(a,b) + \beta)}{\mathrm{Beta}(\beta, \beta)} & \text{otherwise.} \end{cases}$$

$m_a = \sum_{j \neq i} \mathbf{Z}(j, a)$ is the size of the ath functional group disregarding the assignment of the ith node. This posterior likelihood can be evaluated efficiently since we only need to compute $\mathbf{M}_+^{(n)}$ and $\mathbf{M}_-^{(n)}$ and evaluate the Beta function for entries affected by the considered assignment change.

3   Results and Discussion

3.1   Group Membership Visualization

The Gibbs samling result in a posterior distribution of group assignments which makes visualization hard. Thus, here we use the Evidence Accumulation Clustering (EAC) framework [5] to summarize and visualize the MAP solutions from $r = 100$ runs, each ran for 1000 iterations and with α and β fixed to 1. From the MAP solutions we generated the voxel by voxel co-occurrence matrix $\mathbf{C} = \frac{1}{r}\sum_{i=1}^{r}(\mathbf{Z}^{(i)}\mathbf{Z}^{(i)\top} - \mathbf{I})$ where $\mathbf{C}(i,j)$ is the empirical probability that voxels i and j were observed in the same group. Using $\mathbf{C}$ agglomerative hierarchical clustering based on average linkage was performed. We compare the clustering found by IRM with a simpler approach where the mean adjacency matrix $\mathbf{S} = \frac{1}{N}\sum_{n=1}^{N}\mathbf{A}^{(n)}$ was used for agglomerative hierarchical clustering in place of $\mathbf{C}$.

The median number of groups found in the 100 runs was 16 (range 14-18). The average normalized mutual information (NMI) between each pair of the MAP solutions was 0.78 (std=0.04) and all pairs were highly significant ($p < 0.001$, as tested using permutation testing) indicating that the groupings found by IRM are stable across runs. The cophenetic correlation can be used to assess the dispersion of the co-occurrence matrix $\mathbf{C}$ and is therefore also a measure of the stability of IRM solutions across random initializations[2]. The cophenetic correlation was 0.98.

The left part of Fig. 1 shows the grouping found by EAC of the IRM MAP solutions. The voxel groups are rendered on the anatomical mask used (gray: putamen; purple: caudate nucleus) and shown next to the dendrogram. The groups are symmetric between hemispheres, i.e. same sub-territory in left and right striatum are grouped together, suggesting that the IRM is able to group voxels in bilateral hemispheres which are involved in the same neural computations. Using the hierarchical clustering one can assess relations between groups. Again, the model is able to extract meaningful anatomical information, since here the green part of the tree defines the putamen, the red part defines the ventral part of caudate head/tail, while the blue part defines the dorsal part of caudate head/tail.

The right part of Fig. 1 show the voxel groups and dendrogram found by average linkage clustering of $\mathbf{S}$. Besides cluster 10, which is a large cluster of both right and left caudate nucleus, the clustering of $\mathbf{S}$ does not show the symmetry of the IRM grouping but here the groups are in general lateralized to either left or right striatum and are in general groups of nearby voxels. The green part of the dendrogram defines right posterior putamen, blue is bilateral caudate nucleus and anterior putamen, while red reflects posterior putamen. The cophenetic correlation of the hierarchical clustering based on $\mathbf{S}$ was 0.67 indicating the dendrogram is representing less of the information in the data compared to the dendrogram of the IRM model.
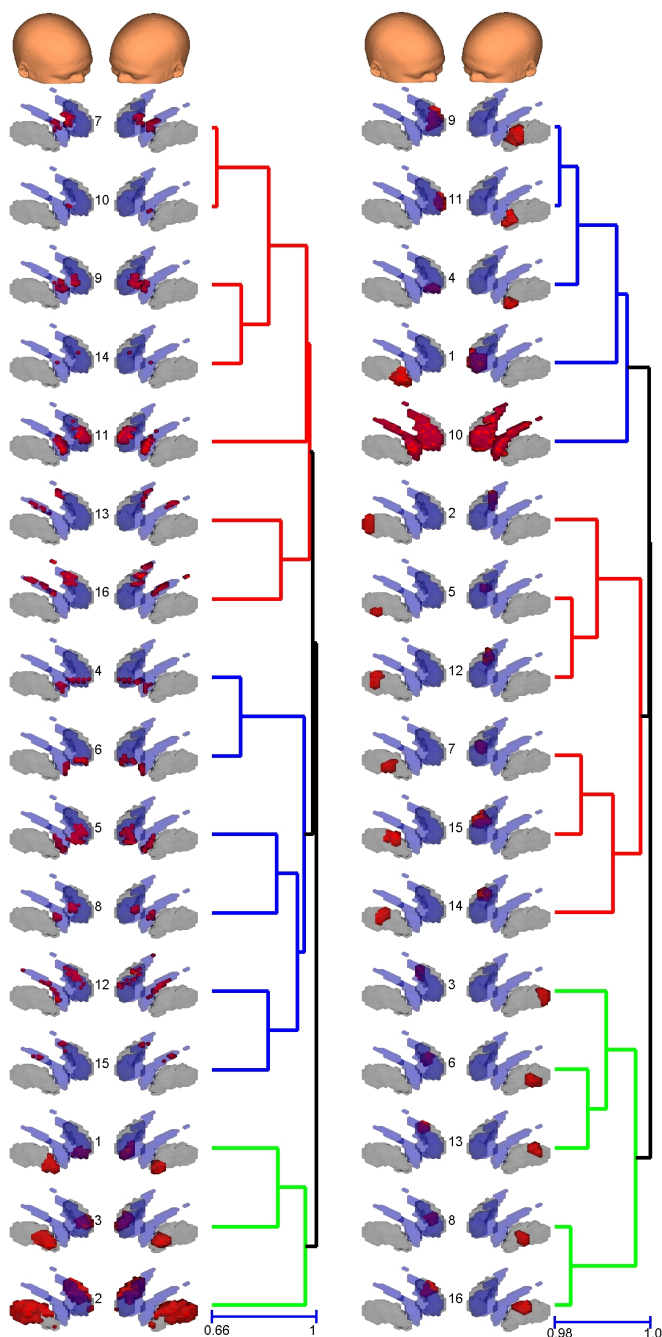
Fig. 1. Left; agglomerative hierarchical clustering of the co-occurence matrix **C** of the MAP solutions found from 100 starts of the IRM inference. Right; for comparison a simpler approach where the average adjacency matrix **S** was used for agglomerative hierarchical clustering in place of **C**.

3.2   Varying Hyperparameters

To test the reproducibility of groupings found for various choices of hyperparameters we split the subject sample in half and ran the IRM on each subsample and calculated the mutual information (MI) and normalized mutual information (NMI) between the MAP solutions found from each subsample. This was repeated for 10 different splits for each hyperparameters value. Fig. 2(a) show the mean (std) log likelihood, MI, NMI and number of groups found when varying $\log_{10}(\alpha)$ from -15 to 15 keeping $\beta = 1$. Likewise, Fig. 2(b) shows the same when varying $\log_{10}(\beta)$ from -6 to 2 keeping $\alpha = 1$. The IRM is seen to be very robust for the choice of α, which controls the prior belief on the group distributions, where both the MI and NMI remain constant over the wide α-range investigated. The choice of β, which controls the prior belief on group link probabilities, have a stronger influence of the MI and NMI. Here, NMI and log likelihood peaks at $\beta = 0.1$. At $\beta = 100$ perfect repetition is found between the two splits (NMI=1) although only 2 groups are found, so the information maintained in this grouping is low as reflected by the low MI.

(a) Varying α (b) Varying β
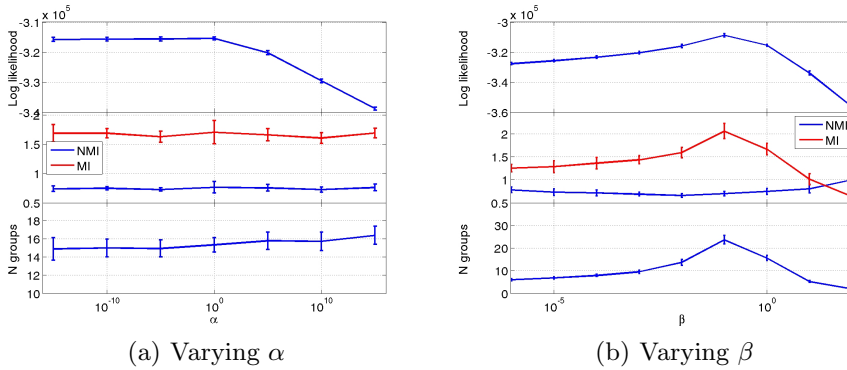
Fig. 2. Log likelihood (top panel), mutual information (MI) and normalized mutual information (NMI) (middle panel) and number of groups (lower panel) between the MAP solutions found by randomly splitting the subjects in two half. Mean and standard deviations are shown for 10 splits for each α (a) and β (b). The IRM is seen to be very robust to the choice of α where both the MI and NMI remain constant over the wide α-range investigated. The choice of β have a stronger influence of the MI and NMI. Here, NMI and log likelihood peaks at $\beta = 0.1$. At $\beta = 100$ perfect repetition is found between the two splits (NMI=1) although here the number of groups found is 2, so the information maintained in this grouping is low as reflected by the low MI

3.3   Link Prediction

To evaluate the model's ability to predict unobserved links we treated at random 2.5% of the links and an equivalent number of non-links in each adjacency matrix as missing. The area under the curve (AUC) of the receiver-operator characteristics was used as performance measure to evaluate how well the model
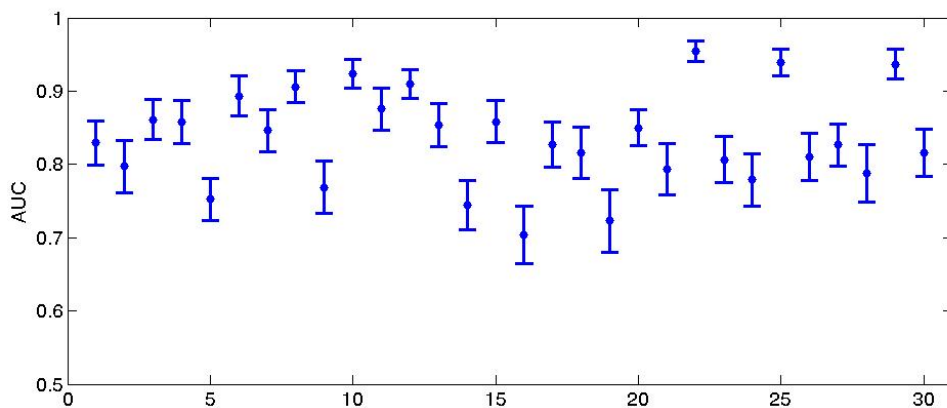
Fig. 3. The area under the curve (AUC) of the receiver-operator characteristics for the link prediction for each of the 30 subjects. The mean (std) across subjects were 0.83 (0.06).

was able to predict these unobserved links. Fig. 3 shows the mean (std) AUC for each of the 30 subjects of 100 model runs. Across subjects the mean (std) AUC was found to be 0.83 (0.06) which is well above chance for all subjects.

4   Conclusion

In this work we used the Infinite Relational Model to infer functional groupings in the human striatum. We show that the groups found are symmetric between hemispheres indicating that the IRM is able to find groups of voxels which are involved in the same neural computations. We evaluate the model's reproducibility by splitting the subject sample in half and compute mutual information and normalized mutual information between splits when varying the hyperparameters of the model. We show that the solutions are very robust to the choice of α, which controls the grouping distribution, while the solutions are more sensitive to the choice of β, which controls the prior belief in group link probabilities. Further, we show that the model is able to predict missing links in the graph well above chance.

References

1. Alexander, G.E., Crutcher, M.D., DeLong, M.R.: Basal ganglia-thalamocortical circuits: Parallel substrates for motor, oculomotor, "prefrontal" and "limbic" functions. Progress in Brain Research 85, 119–146 (1991)
2. Brunet, J.P., Tamayo, P., Golub, T.R., Mesirov, J.P.: Metagenes and molecular pattern discovery using matrix factorization. Proceedings of the National Academy of Sciences of the United States of America 101(12), 4164–4169 (2004)

3. Bullmore, E.T., Bassett, D.S.: Brain graphs: graphical models of the human brain connectome. Annual Review of Clinical Psychology 7, 113–140 (2011)
4. Doyon, J., Bellec, P., Amsel, R., Penhune, V., Monchi, O., Carrier, J., Lehéricy, S., Benali, H.: Contributions of the basal ganglia and functionally related brain structures to motor learning. Behavioural Brain Research 199(1), 61–75 (2009)
5. Fred, A.L.N., Jain, A.K.: Combining multiple clusterings using evidence accumulation. IEEE Transactions on Pattern Analysis and Machine Intelligence 27(6), 835–850 (2005)
6. Haber, S.: The primate basal ganglia: parallel and integrative networks. Journal of Chemical Neuroanatomy 26(4), 317–330 (2003)
7. Jain, S., Neal, R.M.: A Split-Merge Markov chain Monte Carlo Procedure for the Dirichlet Process Mixture Model. Journal of Computational and Graphical Statistics 13(1), 158–182 (2004)
8. Kemp, C., Tenenbaum, J., Griffiths, T., Yamada, T., Ueda, N.: Learning systems of concepts with an infinite relational model. In: Proceedings of the National Conference on Artificial Intelligence, vol. 21, pp. 381–388. AAAI Press, MIT Press, Menlo Park, Cambridge (1999, 2006)
9. Lancaster, J.L., Woldorff, M.G., Parsons, L.M., Liotti, M., Freitas, C.S., Rainey, L., Kochunov, P.V., Nickerson, D., Mikiten, S.A., Fox, P.T.: Automated Talairach atlas labels for functional brain mapping. Human Brain Mapping 10(3), 120–131 (2000)
10. Maldjian, J.A., Laurienti, P.J., Kraft, R.A., Burdette, J.H.: An automated method for neuroanatomic and cytoarchitectonic atlas-based interrogation of fMRI data sets. NeuroImage 19(3), 1233–1239 (2003)
11. Mehler-Wex, C., Riederer, P., Gerlach, M.: Dopaminergic dysbalance in distinct basal ganglia neurocircuits: implications for the pathophysiology of parkinson's disease, schizophrenia and attention deficit hyperactivity disorder. Neurotoxicity Research 10, 167–179 (2006)
12. Middleton, F.A., Strick, P.L.: Basal-ganglia 'projections' to the prefrontal cortex of the primate. Cerebral Cortex 12(9), 926–935 (2002)
13. Mørup, M., Madsen, K., Dogonowski, A.M., Siebner, H., Hansen, L.: Infinite relational modeling of functional connectivity in resting state fMRI. In: Advances in Neural Information Processing Systems 23, pp. 1750–1758 (2010)
14. Obeso, J.A., Rodríguez-Oroz, M.C., Benitez-Temino, B., Blesa, F.J., Guridi, J., Marin, C., Rodriguez, M.: Functional organization of the basal ganglia: therapeutic implications for Parkinson's disease. Movement Disorders: Official Journal of the Movement Disorder Society 23(suppl. 3), S548–S559 (2008)
15. Sporns, O.: The human connectome: a complex network. Annals of the New York Academy of Sciences 1224(1), 109–125 (2011)
16. Xia, M., He, Y.: Magnetic Resonance Imaging and Graph Theoretical Analysis of Complex Brain Networks in Neuropsychiatric Disorders. Brain Connectivity 1(5), 349–365 (2011)
17. Xu, Z., Tresp, V., Yu, K., Kriegel, H.: Infinite hidden relational models. In: Proceedings of the 22nd International Conference on Uncertainity in Artificial Intelligence (2006)

Coordinated Activation in the Resting Brain is Community Structured

Coordinated activation in the resting brain is community structured

Kasper Winther Andersen[a,b,*], Kristoffer H. Madsen[b], Hartwig Roman Siebner[b,c,d], Mikkel N. Schmidt[a], Morten Mørup[a], Lars Kai Hansen[a]

[a]*Department of Applied Mathematics and Computer Science, Technical University of Denmark, Matematiktorvet, Bygning 303 B, 2800 Kgs. Lyngby, Denmark*
[b]*Danish Research Centre for Magnetic Resonance, Centre for Functional and Diagnostic Imaging and Research, Copenhagen University Hospital Hvidovre, Kettegaard Alle 30, 2650 Hvidovre, Denmark*
[c]*Department of Clinical Medicine, Faculty of Health and Medical Sciences, University of Copenhagen, Blegdamsvej 3B, 2200 København N, Denmark*
[d]*Department of Neurology, Copenhagen University Hospital Bispebjerg, Bispebjerg Bakke 23, 2400 København NV, Denmark*

Abstract

Modeling of resting state functional magnetic resonance imaging (rs-fMRI) data using network models is of increasing interest. Studies indicate that rs-fMRI brain networks are modular and that the global modularity index is linked with clinical and behavioral measures. At an intermediate network level, however, it remains unclear which representation is most suitable for modeling the modular structure in rs-fMRI graphs. In this study we consider three different nonparametric Bayesian models for node clustering in com-

---

[*]Corresponding author
Kasper Winther Andersen, kwjo@dtu.dk, Danish Research Centre for Magnetic Resonance, Department 714, Centre for Functional and Diagnostic Imaging and Research, Copenhagen University Hospital Hvidovre, Kettegaard Alle 30, DK-2650 Hvidovre, Denmark, Tel: (+45) 3862 2852

Email addresses: `kwjo@dtu.dk` (Kasper Winther Andersen), `stoffer@drcmr.dk` (Kristoffer H. Madsen), `hartwig.siebner@drcmr.dk` (Hartwig Roman Siebner), `mns@dtu.dk` (Mikkel N. Schmidt), `mm@dtu.dk` (Morten Mørup), `lkai@dtu.dk` (Lars Kai Hansen)

plex networks. In particular we test their ability to predict unseen network data and the reproducibility of their clusterings for different datasets. The three models considered are the Infinite Relational Model (IRM), Bayesian Community Detection (BCD), and the Infinite Diagonal Model (IDM). The IRM is the most flexible of the three models with no restrictions on the between-cluster link probabilities, the BCD model in contrast restricts the between-cluster link probabilities to be strictly lower than within-cluster link probabilities to conform with the modular community structure typically seen in social networks, while the IDM restricts the individual between-cluster link probabilities to be identical. Using three different datasets comprising healthy volunteers' rs-fMRI we found that the BCD model was in general the most predictive and reproducible of the three. This suggests that rs-fMRI data exhibits community structure and furthermore points to the significance of modeling heterogeneous between-cluster link probabilities.

Highlights

1. Three nonparametric Bayesian models for node clustering are used to model rs-fMRI.

2. Models predictability and reproducibility are extensively evaluated using resampling.

3. The community structure model shows better predictability and reproducibility.

4. This finding suggests that rs-fMRI graphs exhibit community structure.

5. Modelling between-cluster link probabilities adds important information.

1. Introduction

Analysis of resting state functional magnetic resonance imaging (rs-fMRI) has emerged as a powerful research tool to study whole-brain functional connectivity. Since rs-fMRI provides information about intrinsic fluctuations in functional connectivity within and among brain networks, many conventional analysis schemes applied in task-related fMRI studies are irrelevant. Hence, a number of new techniques have been developed based on identification of stable spatio-temporal multivariate structure in the brain wide set of blood oxygen level dependent (BOLD) time series.

Using correlation methods or flexible multivariate techniques like independent component analysis (ICA) it has been shown that the BOLD signals of distant brain regions are coordinated suggesting interaction as they form so-called resting-state networks. The number and precise definition of these networks are controversial but several networks are broadly accepted, including the default mode network, motor network, visual network, fronto-parietal, dorsal attention network (Damoiseaux et al., 2006).

Complex network analysis is a very active research field (Barabási, 2003) that has already found application in neuroimaging and in modeling resting state connectivity (Bullmore and Bassett, 2011; Sporns, 2011). The basic object is the 'network graph'. When applied to neuroimage analysis the network graph is formed by brain regions represented as nodes. Nodes are connected by a link if brain regions are co-activated above a certain threshold.

3

In rs-fMRI co-activation is often measured simply by calculating correlation between time series.

Network structure can be studied at many levels, from local motifs to global features like scale free link distributions signifying long-range coordination (van den Heuvel et al., 2008). At the intermediate level we may identify clusters of highly linked nodes, i.e., high within-cluster link density and low link density to nodes in other clusters. By analogy to social networks such groups are referred to as *communities*. The presence of community structure in a network can be quantified by the global *modularity* index (Newman, 2006). Modularity can also be used to identify communities, i.e., by clustering nodes such that the modularity index is maximized (Newman, 2006; Lehmann and Hansen, 2007). Modularity has been shown to be related to motor learning (Bassett et al., 2011), visual working memory (Stevens et al., 2012), and Meunier et al. (2009) found that modularity is reduced during normal aging. Evidence is emerging that global modularity can be used as a bio-marker. For instance patients with childhood-onset schizophrenia have reduced modularity of their resting state networks (Alexander-Bloch et al., 2010). However, focusing on modularity as the single summary of a complex network may be overly simplistic as the modularity measure does not account for variability in the inter-linking relations between functional clusters. Hence, modularity driven clustering might not reveal all salient aspects of community structure in a network. Indeed, modularity has been criticized for its lack of flexibility as a measure of community structure (Fortunato and Barthélemy, 2007).

A better understanding of this important mid-level structure in brain

networks requires methods that can capture more informative representations of community structure in brain networks. For this we turn to a family of expressive generative network models. Relational Models are statistical generalizations of graph clustering that consider not only the within-cluster density but also take the specific relations between clusters into consideration. The Infinite Relational Model (IRM) (Kemp et al., 2006; Xu et al., 2006) is a non-parametric generalization of the stochastic block model (Nowicki and Snijders, 2001), for inference of such generalized group structure in complex networks. As the IRM representation considers both linking within and between groups, a highly inter-linked group of nodes could in fact be clustered in different groups if they link in different ways to other clusters, i.e., the IRM can infer more general group structures beyond the conventional community structure. An additional feature of the IRM type of model is that it conveniently allows for analysis of multi-graph networks, which for neuroimaging data could be graphs from multiple sessions or subjects. For multi subject analysis one could look for a common node clustering structure over subjects but allow individual subject cluster linking densities (Mørup et al., 2010) or test the hypothesis that both clustering and link structure are shared between all subjects (Andersen et al., 2012b).

A constrained variant of the IRM representing the community structure of graphs in the sense of grouping highly connected node sets was proposed recently in (Mørup and Schmidt, 2012). The Bayesian Community Detection (BCD) scheme restricts the between-cluster link densities to be strictly lower than within-cluster link densities, thus constraining the more general IRM to conform with the notion of a community in a social network. An additional

constraint can be introduced by considering the so-called Infinite Diagonal Model (IDM) (Mørup and Schmidt, 2012; Schmidt and Mørup, 2013). The IDM allows for differential within-cluster link densities but share a common between-cluster density and as such the variability in the link densities between clusters is neglected when inferring the clustering structure. Thus, the IDM can be considered a probabilistic representation of the modularity concept.

It should be noted that basing the graph on simple time series correlation certain metrical properties can be expected. In particular, if a node A is highly correlated with node B, and B is highly correlated with C, then there is a lower limit on the correlation between nodes A and C which can be inferred by the triangle inequality (Zalesky et al., 2012). This bound will support the formation of community structure, as in social relations: 'Friends of friends are friends', however, we also note that by thresholding the correlation, the impact on the community structure of these geometrical constraints is non-trivial.

Spatial grouping of brain regions by similarity of BOLD time series as pursued in the present work can be seen as complementary to classical approaches to spatial grouping such as time series clustering (Goutte et al., 1999) and independent component analysis (ICA) (McKeown et al., 1998; McKeown, 2003). Compared with conventional clustering, the relational modeling approach has the advantage that clusters are formed by considering the connectivity patterns both within and between clusters, and furthermore relational models avoid the formation of a group prototype, hence allows for more flexible group structures to be found (Kemp et al., 2006). The use of

ICA is based on strong assumptions of independency, which can be questioned in the resting state as it has been observed that groups are negatively correlated in time (Fox et al., 2005).

In this study, we apply the above-mentioned community detection schemes to rs-fMRI data acquired in three cohorts of healthy volunteers and investigate to which degree functional brain networks as measured by rs-fMRI exhibit community structure. The three Bayesian relational methods, i.e. IRM, BCD, and IDM, for inference of group structure in complex networks differ only in the way the link probabilities between clusters are modeled. The rich link structures of the relational models can be seen as a way of inferring functional integration at the inter-community level as discussed in (Hagmann et al., 2008; Sporns, 2013). We evaluate the performance of these models with respect to their ability to predict out-of-sample data (predictability) and the robustness of their clustering under re-sampling of data (reproducibility) using the NPAIRS split-half framework (Strother et al., 2002). The evaluation is carried out on three datasets from different sites and the models are evaluated both within and between sites for several thresholds of the correlation matrices. The work in this paper builds on work presented in Andersen et al. (2012b).

2. Methods

For generality we investigate three rs-fMRI datasets. One dataset acquired locally at the Danish Research Centre for Magnetic Resonance (Copenhagen) and two other rs-fMRI datasets publicly available in the FCON1000 database (Biswal et al., 2010) (viz., the 'Beijing' and the 'Leipzig' datasets).

2.1. Copenhagen data

The Copenhagen dataset included thirty healthy controls with no history of neurological or psychiatric disease. At the day of scanning all subjects were asked to refrain from caffeine, cigarettes or alcohol intake at least six hours prior to the scanning session. All subjects gave written informed consent prior to scanning and the study was approved by the local scientific ethics committee of Copenhagen and Frederiksberg Communities (protocol no. KF01 - 131/03 with addendum). The Edinburgh handedness inventory (Oldfield, 1971) revealed that all participants except two were right handed.

All MRI measurements were performed on a 3.0 Tesla Magnetom Trio scanner (Simens, Erlangen, Germany). Each participant underwent an MRI session including a structural scan as well as a functional scan during rest both with full brain coverage. During the functional scan subjects were instructed to rest with their eyes closed without falling asleep, and refrain from any voluntary motor or cognitive activity.

The first scan during each session was the rs-fMRI functional scan which consisted of a T2* weigthed echo planar imaging (EPI) sequence with a repetition time of 2490 ms, echo time 30 ms and flip angle 90 degrees. Over a period of 20 minutes we collected 480 brain volumes each consisting of 42 axial slices with an isotropic resolution of 3 mm, field of view (FOV): 192x192 mm. During scanning we monitored the subjects cardiac and respiratory cycles using an infrared pulse oximeter and a pneumatic thoracic belt. The structural scan was based on a magnetization prepared rapid gradient echo (MPRAGE) sequence with the following parameters: Repetition time (TR) = 1550 ms, echo time (TE) = 3.04, inversion time (IT) = 800 ms; 192 sagittal

slices; 1 mm isotropic resolution; FOV = 256 mm; flip-angle = 9 degrees.

The functional images were preprocessed using statistical parametric mapping software (SPM8, Wellcome Trust Centre for Neuroimaging, `http://www.fil.ion.ucl.ac.uk/spm`) implemented in Matlab 7.9 (MathWorks, Massachusetts, USA). In order to allow stabilization of T1 equilibrium effects we discarded the first five volumes prior to analysis. The remaining 475 brain volumes were realigned to the time-series mean using a two-step procedure and then co-registered to the same-session T1-weighted MPRAGE scan by a 6-parameter rigid-body transformation. The T1-weighted scan was spatially normalised to the Montreal Neurological Institute (MNI) 305 standard template using the unified segmentation/normalisation procedure as implemented in SPM8 (Ashburner and Friston, 2005). Subsequently the same normalisation parameters were used to normalise the EPI images.

Both hardware instability and unwanted physiological effects (such as cardiac pulsation and respiration) produce signal changes in fMRI time-series (Smith et al., 1999; Dagli et al., 1999; Glover et al., 2000; Lund et al., 2006). These signal changes may give rise to signal fluctuation resembling those typically observed in rs-fMRI data (Birn et al., 2006). In order to reduce these effects prior to extraction of time series for the regions of interest we applied comprehensive temporal filtering of cardiac, respiratory and motion related effects. The filter included a high-pass filter based on discrete cosine basis functions (cut-off frequency 1/128 Hz). Cardiac and respiratory modeling cycles were modeled using Fourier expansions of the aliased cardiac (10 parameters) and respiratory (6 parameters) cycles as well as first order cardiac by respiration cycles interaction (4 parameters) effects (Glover et al.,

2000). Residual motion effects (24 parameters) were modeled using a Taylor expansion of the estimated movement parameters including spin-history effects (Friston et al., 1996). Changes in the respiration volume over time has been demonstrated to produce signal changes resembling those observed in rs-fMRI (Birn et al., 2006) we model these changes by included 41 time delayed versions (time delay between 20 and -20 seconds in one second intervals) of the respiration volume. Finally the filter included individual time series from cerebrospinal fluid voxels and white matter voxels from both the right and left hemispheres.

2.2. Beijing and Leipzig data

Two other datasets were used from the FCON1000 database (Biswal et al., 2010) (`http://fcon_1000.projects.nitrc.org`). See Appendix A for a list of subjects used. The Beijing dataset consist of 42 of the subjects from the Beijing Zang set. The dataset is recorded with 33 slices using TR=2000ms and with 225 brain volumes. The Leipzig dataset consist of 37 subjects (21 females), ages 20-42, TR=2300ms, 34 axial slices, and 195 brain volumes.

For both datasets the first 5 volumes have already been discarded. Preprocessing was done in SPM8 and included realigning to time-series mean for motion correction and normalising to standard MNI space using the template EPI image included in SPM.

2.3. Graph construction

We extracted the mean signal in each of the 116 regions covered in the AAL database (Tzourio-Mazoyer et al., 2002) and constructed the correlation

matrix for each subject. Since this matrix is symmetric only the upper diagonal is further considered. Each subject's correlation matrix were binarized at an individual level to maintain the highest d-percent correlations corresponding to having a graph link density at d-percent. After thresholding an adjacency matrix $\mathbf{A}$ is retrieved where $A_{i,j}$ is the (i,j)-th element of $\mathbf{A}$ and $A_{i,j} = 1$ if there is a link between nodes i and j and $A_{i,j} = 0$ otherwise. Since we model multiple subjects $\mathbf{A}^{(n)}$ denotes the adjacency matrix corresponding to subject n.

2.4. The models

This section will provide an overview of the models considered in this paper. For a more in depth description please refer to Schmidt and Mørup (2013). The goal is to group nodes into non-overlapping clusters, such that a common node-clustering across subjects is retrieved. Let $\mathbf{z}$ be the vector of nodes assignments where z_i takes the integer value corresponding to the cluster index node i belongs to. The models used are all generative models meaning that given the model definition and the model parameters one can generate new graphs by drawing samples from the model. The models differ in the way they model the link probability between and within clusters. Let $\rho_{k,l}$ represent the link probability between clusters k and l. Since we here consider undirected graphs $\boldsymbol{\rho}$ is symmetric.

2.4.1. The Infinite Relational Model

In IRM link probabilities within and between clusters are modeled individually and without restrictions. As such the model allows for complex relations between clusters, and thus allow for flexible clustering of nodes.

11

Consider generating graphs from this model. The first step is to draw assignments of nodes into clusters, which is done using the Chinese Restaurant Process (CRP) (Aldous, 1985) using the hyper-parameter α. The CRP generates a cluster assignment, where α controls the number of clusters generated, where larger α will generate more clusters. Next, the link probabilities within and between clusters $\rho_{k,l}$ are generated from the symmetric Beta distribution using the hyper-parameter β. Finally, the cluster assignments and the link densities are used to generate links between nodes. This is done using the Bernoulli distribution, where the probability of a link (success) between nodes i and j is determined by the clusters (z_i and z_j) the nodes belong to.

The generative model can be summarized as:

Infinite Relational Model	
Cluster assignments:	$\mathbf{z} \sim \text{CRP}(\alpha)$
Link probabilities:	$\rho_{k,l} \sim \text{Beta}(\beta, \beta)$
Links:	$A_{i,j}^{(n)} \sim \text{Bernoulli}(\rho_{z_i,z_j})$

In Appendix B.1 we derive the likelihood function for the IRM which is used in model inference.

2.4.2. Infinite Diagonal Model

The model termed Infinite Diagonal Model (IDM) (Mørup and Schmidt, 2012) is a special case of the IRM where link probabilities between clusters are constrained to be equal. As such, the IDM does not model the relation between clusters but has a constant background link probability. The only

difference in terms of the model formulation is then

$$\rho_{k,l} = \begin{cases} \rho_{k,k} & \text{if } k = l \\ \rho_b & \text{otherwise.} \end{cases}$$

2.4.3. Bayesian Community Detection

A network community is defined as a group of nodes with more dense linking internally than externally. The Bayesian Community Detection (BCD) model proposed in (Mørup and Schmidt, 2012) enforces larger within-cluster link probabilities than between-cluster link probabilities. Like IRM, the cluster assignments are first generated using the CRP. A cluster-gap is then drawn from a symmetric Beta distribution with hyperparameter v. The within-cluster link probabilities are then drawn for each cluster again using the Beta distribution. The between-cluster link probabilities are subsequently drawn using the incomplete Beta distribution BetaInc(a, b, x) constrained to the interval $[0, x]$, with the density, $p(\theta) = \frac{1}{B_x(a,b)} \theta^{a-1} (1-\theta)^{b-1}$, where $B_x(a, b)$ is the incomplete beta function. The between-cluster link probability between two clusters k and l can then at most be as high as the smallest of the two within-cluster link probabilities multiplied by the cluster gap. The lower the gap-value the higher difference in within and between-cluster link probability. Finally, links are drawn using the Bernoulli distribution just like the other models. The generative model for BCD can thus be summarized as:

13

Bayesian Community Detection

$$\text{Cluster assignments}: \quad \mathbf{z} \sim \text{CRP}(\alpha)$$

$$\text{Cluster gap}: \quad \gamma \sim \text{Beta}(v, v)$$

$$\text{Link probability}: \quad \rho_{k,l} \sim \begin{cases} \text{Beta}(\beta, \beta) & \text{if } k = l \\ \text{BetaInc}(\beta, \beta, w_{k,l}) & \text{otherwise.} \end{cases}$$

$$\text{where } w_{k,l} = \min[\gamma \rho_{ll}, \gamma \rho_{kk}]$$

$$\text{Links}: \quad A_{i,j}^{(n)} \sim \text{Bernoulli}(\rho_{z_i, z_j})$$

2.5. Example 1

For illustration we generate a graph consisting of 50 nodes from each of the three models with $\alpha = 5, \beta = 1$. For the BCD model we set $v = 1$. Figure 1 shows the generated graphs. The plots are a combination of both the cluster assignment matrix, the adjacency matrix, and the link probability matrix. The adjacency matrix $\mathbf{A}$ is plotted, where links between nodes are indicated by small black dots. Cluster membership is indicated with the colors to the left and top of the adjacency matrix and the link probability matrix is indicated with gray shaded background. For IRM there are no restrictions in the link probability values, resulting in some between-cluster link probabilities being larger than within-cluster link probabilities. For the BCD model the between-cluster link probability between two clusters are restricted to be smaller than the within-cluster link probability times the gap. The gap was drawn from the Beta distribution and in this case the gap is $\gamma = 0.96$. For the IDM model all the between-cluster link probabilities are

14
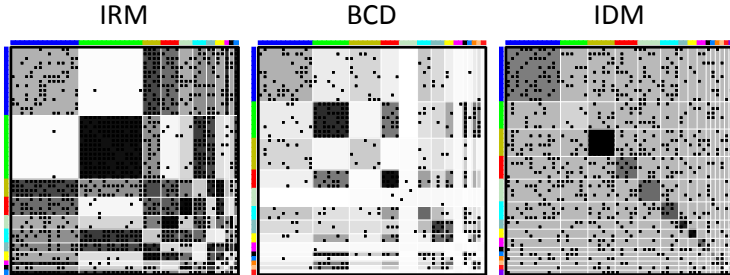
Figure 1: Example 1. Figure illustrating data drawn from each of the three models, IRM, BCD, and IDM respectively. Plots illustrates both the adjacency matrix **A** (links indicated by small black squared dots), cluster membership **z** as color codes to the left and top of the adjacency matrix, and link probability matrix $\boldsymbol{\rho}$ as gray shading of the matrix elements.

equal meaning that clusters are only defined in the way they link internally in the clusters.

2.6. IRM and IDM model inference

In the previous sections we defined the generative models, which allow one to generate data by sampling from the model. However, we are interested in inferring the model parameters given the data. By using the model definition the joint likelihood can be written and by using Bayes theorem an expression for the posterior distribution can be found. It is then possible to sample from this posterior distribution using Markov chain Monte Carlo sampling (MCMC) methods. For IRM and IDM the link probabilities can analytically be integrated out which means that we only have to sample over the node assignments. For that Gibbs sampling in combination with split-merge Metropolis-Hastings updates (Jain and Neal, 2004; Kemp et al., 2006; Mørup et al., 2010) is used. Below is a description of these two steps.

15

Gibbs sampling is a Markov-chain Monte Carlo sampling method. For each scan of the Gibbs sampler each node's cluster assignment is updated using the conditional distribution of that node's assignment given the assignments of the remaining nodes. For IRM the conditional distribution is derived in Appendix B.1 (equation B.2) and for IDM it is derived in Appendix B.2 (equation B.3).

Given the incremental nature of the Gibbs sampling algorithm it has difficulties escaping local maxima of the probability landscape. E.g. it is hard to split a single cluster into two new clusters since this requires that nodes are moved one at a time from a cluster to the other cluster. To overcome this we use a restricted split-merge sampling scheme (Jain and Neal, 2004), which potentially merges two existing clusters into one or split a single cluster into two clusters. At each step of the algorithm two nodes are selected at random with uniform probability. If the two selected nodes currently are assigned different clusters then an assignment is proposed where these two clusters are merged into one cluster. On the contrary, if these two selected nodes are currently assigned to the same cluster then a new assignment is proposed where all nodes assigned to this cluster are split into two separate clusters. The split-proposal is found using a restricted Gibbs-sampling procedure. First a launch state is found by allocating the two nodes to two different empty clusters as proposed in (Dahl, 2005). Then remaining nodes are in random order assigned to either of the two clusters based upon their conditional probability. This state is then referred to as the launch state. The launch state is refined by restricted Gibbs sampling steps where nodes from the two new clusters can be re-assigned either of the two

16

clusters based on the conditional probability (equation B.2 and B.3). This procedure is restricted because only nodes from the cluster from which the nodes originally came from are re-assigned and they can only be assigned to either of the two new clusters. The proposed configuration is then sampled from the launch state. If this proposed state in the Markov chain is accepted with the Metropolis-Hasting acceptance probability then this becomes the new state else the old state is kept as the new state.

2.7. BCD model inference

In IRM and IDM we are able to marginalize link probabilities ($\boldsymbol{\rho}$) out. This is not the case in BCD because between-cluster link probabilities are dependent of the within-cluster link probabilities. However, the vast majority of the parameters, namely the between-cluster link probabilities, can be integrated out (Appendix B.3). The remaining parameters $\mathbf{z}, \dot{\boldsymbol{\rho}}$, and γ are sampled using MCMC, where $\dot{\boldsymbol{\rho}}$ refer to the within-cluster link probabilities (the diagonal of $\boldsymbol{\rho}$). The within-cluster link probabilities and cluster gaps are sampled with Metropolis-Hastings. The cluster assignments $\mathbf{z}$ are like the IRM sampled with Gibbs sampling and split-merge moves, however new possible values for the within link probabilities and cluster gaps are first drawn from their prior. In Appendix B.3 we derive the conditional distributions used in the sampling. For further information please see Mørup and Schmidt (2012).

2.8. Example 2

We illustrate differences in cluster assignments and link probabilities inferred by each of the three models. We generate a synthetic graph with

17

40 nodes, 10 nodes in each of four clusters. The example is designed such that *cluster1* and *cluster2* share the same within and between-cluster link probabilities, however only *cluster2* is connected with *cluster3*. *Cluster3* and *cluster4* have low within-cluster probabilities but high between-cluster link probability. *Cluster3* and *cluster4* are not connected to *cluster1* and *cluster2*. The first row in Figure 2 show the true assignment vector ($\mathbf{z}$) coded as a 1-of-n matrix and the true link probabilities. The next rows show the assignments and link probabilities inferred by the IRM, BCD, and IDM models respectively. Except for a single node IRM finds the correct grouping structure. BCD assigns the first two clusters correctly and mislabels the same node as IRM, but BCD has difficulties with the remaining nodes because the true model has higher between-cluster than within-cluster link probabilities. Since IDM does not model the between-cluster link probabilities, it groups the first two clusters together and the next two clusters together.

2.9. NPAIRS Evaluation Criteria

To evaluate the performance of the models, we used the NPAIRS split-half evaluation framework (Strother et al., 2002). Under this framework the set of subjects were split into two half-splits (S1 and S2) and models was inferred on each half-split enabling us to estimate the predictability and reproducibility of the models. The models' predictability was evaluated using test log likelihood. The node assignment and link probabilities from the sample with the highest value of the posterior distribution were used to calculate the test log likelihood of the other split. The test log likelihood was calculated for both splits (with the other split as training data) and the average test log likelihood was calculated and used as the predictability
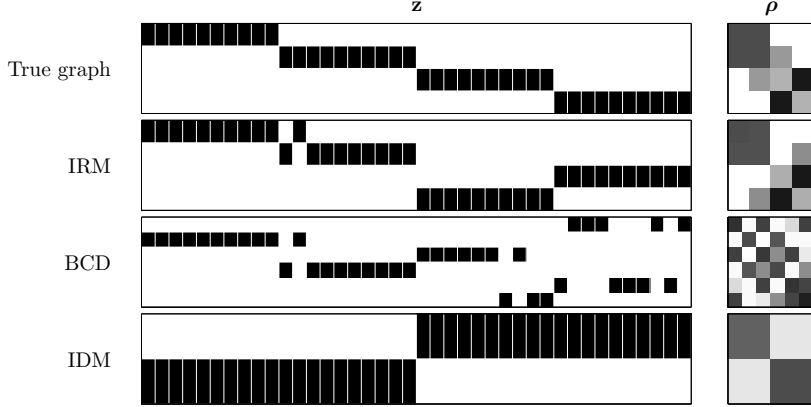
Figure 2: Example 2. First row show the true assignments coded as a 1-of-n matrix and the true link probabilities. The next rows show the structure and link probabilities inferred by IRM, BCD, and IDM respectively.

measure. The test log likelihood for split S2 (using the model parameters inferred using split S1) was calculated by

$$\log P(\mathbf{A}^{S2,(1)}, ..., \mathbf{A}^{S2,(N)}|\boldsymbol{\rho}, \mathbf{z}) =$$

$$\frac{1}{N} \sum_{n=1}^{N} \sum_{j>i} \left[A_{i,j}^{S2,(n)} \log(\rho_{z_i,z_j}) + (1 - A_{i,j}^{S2,(n)}) \log(1 - \rho_{z_i,z_j}) \right]. \quad (1)$$

We measured the reproducibility of the models using normalized mutual information between assignment matrices ($\mathbf{z}^{S1}$ and $\mathbf{z}^{S2}$) of the sample with the highest value of the posterior distribution inferred using the two different splits.

$$\text{NMI} = \frac{2\text{MI}(\mathbf{z}^{S1}, \mathbf{z}^{S2})}{\text{MI}(\mathbf{z}^{S1}, \mathbf{z}^{S1}) + \text{MI}(\mathbf{z}^{S2}, \mathbf{z}^{S2})}, \quad (2)$$

19

where

$$\text{MI}(\mathbf{z}^{\text{S1}}, \mathbf{z}^{\text{S2}}) = \sum_{k=1}^{D1} \sum_{l=1}^{D2} p(z^{\text{S1}} = k, z^{\text{S2}} = l) \log \left(\frac{p(z^{\text{S1}} = k, z^{\text{S2}} = l)}{p(z^{\text{S1}} = k) p(z^{\text{S2}} = l)} \right), \quad (3)$$

where $D1$ and $D2$ are the number of clusters inferred using S1 and S2, respectively.

The model used in e.g. Mørup et al. (2010) used individual subject link probabilities, that is each subject was modeled with her own link probability matrix while sharing the node assignments $\mathbf{z}$. This model allows for subject variability in the communication between clusters and can be used to test for differences in subject populations. However, here we are interested in the models' predictive abilities, that is, how well can a model and its parameters learned from a sub-group of subjects predict the graphs from another group of subjects. Therefore we do not model individual subject link densities but constrain $\boldsymbol{\rho}$ to be common across subjects. The derivation of the models (Appendix B.1) reveal that this amounts to simply summing the adjacency matrices across subjects $\sum_n \mathbf{A}^{(n)} = \mathbf{A}^{\text{tot}}$. This means that under this restricted model definition inference of the latent variables of the model does not scale with the number of graphs (subjects) and therefore our model formulation allows for analysis of large numbers of subjects.

2.10. Experiments

In all the experiments described below nodes were randomly assigned to one of 50 clusters and the sampler ran for 500 iterations. The sample with the highest value of the posterior distribution was then used as representative for a given run. In all experiments $\alpha = \beta = 1$.

2.10.1. Estimated clusters

To inspect the clustering of the different models, the inference procedure were launched 10 times for each model using a graph link density $d = 8\%$ and the sample with the overall highest value of the posterior distribution across the 10 runs was visualized. The reproducibility measured as the mean NMI between the samples with the highest value of the posterior distribution for each run was calculated. Likewise, the clustering similarity between the methods was also estimated by calculating the mean NMI between each pair of the 10 solutions found.

2.10.2. Predictability and reproducibility

We asked how well the clusterings reproduce between datasets and how well the models predict new data. To this end, we evaluated the models using the NPAIRS framework. Subjects were randomly split into two equally sized groups and model inference was conducted on each split. The highest posterior distribution sample was identified for the two splits and NMI between clusterings was calculated as a measure of the models' reproducibility. Using the estimated link probability matrix and assignment from the sample with the highest value of the posterior distribution of one split the test log likelihood for the other split was calculated as a measure of the models' predictability. This was done by for 100 different half-splits of the Copenhagen dataset using 8% graph link density.

2.10.3. Varying link density

For further evaluation of the methods the analysis were repeated within each of the three datasets as well as between the datasets for graph link

densities of $d = 2\%, 4\%, 8\%, 16\%, 32\%$. For analysis done within each individual dataset the subjects were randomly split in half. For the between dataset analysis inference was done within each dataset and NMI and test log likelihood was calculated between datasets. For each link density the log likelihood ratio was calculated as the log likelihood of a random Erdős-Rényi model having the considered link density divided by the log likelihood of the inferred model. This makes the predictability measure more comparable between link densities, however, we note that the log likelihood cannot directly be compared for different link densities as the data itself changes when changing the link densities.

3. Results

3.1. Estimated clusters

Using a graph link density of 8% the reproducibility between solutions found with different restarts was measured as the NMI between the sample with the highest posterior distribution for each run. This was done within all three methods and between methods and results is shown in table 1 along with the number of clusters estimated by each of the methods. For all three methods the clustering for different initializations showed a very high consistency as the NMI was greater than 0.96 for all methods. Also, the number of estimated clusters was very consistent within method, but showed a great between method variability where IRM estimated on average 35.7 clusters, BCD estimated 41.0 and IDM estimated only 18.8. For BCD the mean (std) gap parameter was estimated to 0.88 (0.02). The IRM and BCD clusterings were found to be very similar with a mean NMI of 0.94. The IDM clustering,

22

Method	IRM	BCD	IDM	Mean (std) D
IRM	0.96 (0.01)	-	-	35.7 (1.25)
BCD	0.94 (0.01)	0.96 (0.02)	-	41.0 (2.05)
IDM	0.76 (0.02)	0.75 (0.01)	0.97 (0.02)	18.8 (1.14)

Table 1: The mean(std) of normalized mutual information (NMI) between the clustering of 10 runs within and between method along with the number of clusters (D) estimated with each of the three methods IRM, BCD, and IDM.

however, was less similar to the other two methods with a mean NMI of 0.76 and 0.75 to IRM and BCD respectively.

In figure 3 the samples with the highest value of the posterior distribution across the 10 runs for each method are visualized. The first column shows the link probability matrix $\boldsymbol{\rho}$ which has been permuted such that clusters with the greatest overlap between methods are first. The labels for the clusters can be found in Appendix C. The matrix elements are color-coded in grey-scale according to the value of the link probabilities and the matrix element size indicate the size of the respective cluster. The first 5 clusters were identical between the three methods. The next 12 clusters were identical between IRM and BCD while IDM had all these clusters in one large cluster. When looking at the link probabilities between these 12 clusters it is evident that there is a high link probability within and between these nodes, but subtle differences exist between the different nodes which caused the IRM and BCD to cluster them into separate clusters. Since the IDM model does not consider the between cluster link probabilities these clusters were grouped together in the IDM method. The same were true for the next 6 clusters which were identical for the IRM and BCD and all lumped together in the IDM model

23

because there is a relative high link probability between these clusters. The next three columns show the found clusters in posterior, lateral and superior views of the brain. The clusters are colored according to the colors shown next to the link probability matrices (and the labels given in Appendix C). Brain regions within clusters are connected with lines where line thickness indicates the average link density over subjects for the specific connection. This figure shows that in general the IRM and BCD clusterings were very similar. These two methods produced clusters with relatively few nodes and in general interhemispheric homologues areas were grouped together. IDM also grouped interhemispheric homologues areas together, however as just described this method does not consider specific relations to other brain areas so the clusters produced were larger rather unspecific clusters. For instance the cluster colored in turquoise is a cluster made up of 34 nodes including frontal, occipital, parietal, and temporal lobes.
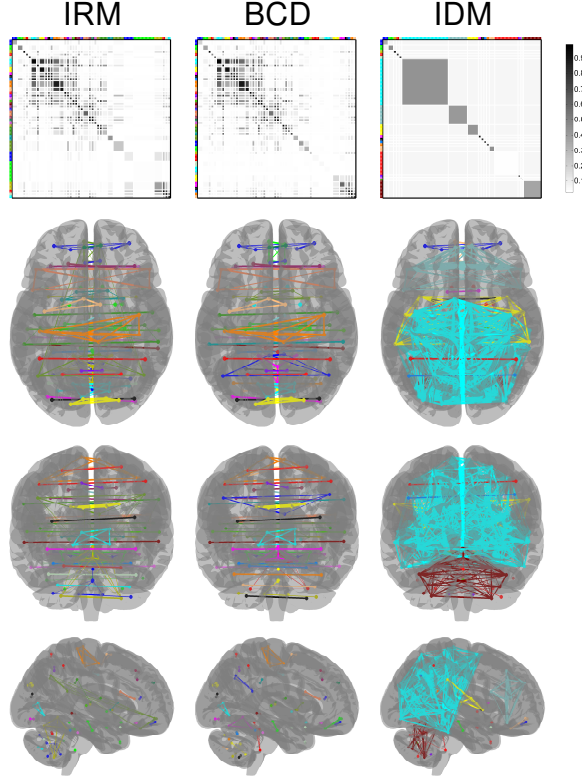
Figure 3: The extracted clusters using the three methods IRM, BCD, and IDM respectively. The first row shows the link probability matrices ρ which has been permuted such that the order of the clusters correspond across methods. The matrix elements are color-coded according to the value of the link probabilities and the size of the matrix elements indicates the size of the respective cluster. The colors next to the matrices correspond to different clusters. The next three rows show the clusters in three different views (superior, posterior, and lateral) of the brain. The clusters are color coded according to the colors next to the link probability matrices and node assignment for each node can be found in Appendix C with the same color as plotted here. Different brain regions within each cluster are connected with lines where the thickness of the line indicates the average link density across subjects for the specific connection.

In figure 4 we show an example of one of the clusters and how the results can be used to inspect connectivity between clusters. A cluster composed of left and right supplementary motor area and left precentral gyrus (A) was selected. This cluster was identical for IRM and BCD while results are not shown for IDM. The figure also displays the 4 clusters with highest between-cluster link probabilities to this cluster. These 4 clusters with highest link probabilities are: (B, $\rho_{A,B} = 0.732$) left and right postcentral gyrus, left and right paracentral lobule and right precentral gyrus; (C, $\rho_{A,C} = 0.714$) left and right middle cingulate gyrus); (D, $\rho_{A,D} = 0.516$); left and right superior frontal gyrus); (E, $\rho_{A,E} = 0.456$) left and right superior temporal gyrus). The line width between clusters in the figure reflects between-cluster link probabilities, likewise the width of the boxes reflect the within-cluster link probabilities.
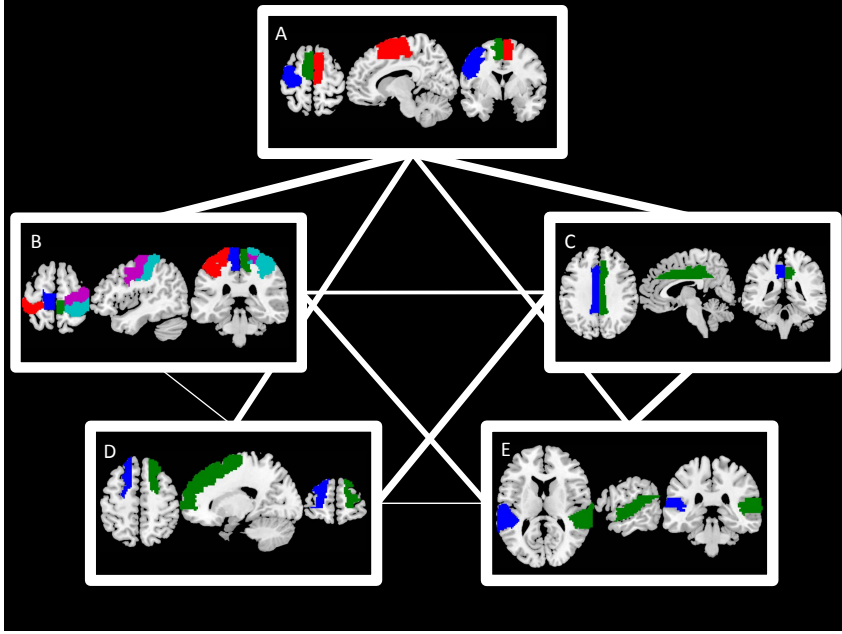
Figure 4: This figure shows a single cluster A composed of the areas left and right supplementary motor area and left precentral gyrus. Also, the 4 clusters with highest between cluster link probability to this cluster is shown. These 4 clusters are (B ($\rho_{A,B} = 0.732$); left and right postcentral gyrus, left and right paracentral lobule and right precentral gyrus), (C ($\rho_{A,C} = 0.714$); left and right middle cingulate gyrus), (D ($\rho_{A,D} = 0.516$); left and right superior frontal gyrus), (E ($\rho_{A,E} = 0.456$); left and right superior temporal gyrus). The line width between nodes reflect the link probabilities between clusters, likewise the width of the boxes reflect the within-cluster link probabilities.

3.2. Predictability and reproducibility

Figure 5 shows the PR scatter plot of the predictability versus reproducibility of the 3 methods using the NPAIRS split-half framework. Clearly, IRM and BCD performed better than IDM in both reproducibility and pre-
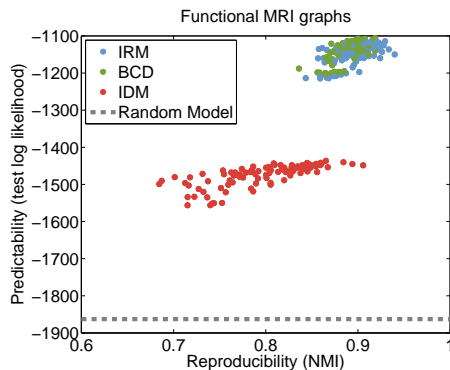
Figure 5: Reproducibility vs predictability plot for the three models using a link density of 8%. IDM and BCD are better in NMI and test log likelihood than IDM ($p < 0.0001$). BCD has better predictability ($p = 0.023$) than IRM while IRM and BCD do not differ in reproducibility ($p = 0.15$).

dictability as measured with NMI and test log likelihood ($p < 0.0001$, permutation test). The IRM and BCD scatter plots overlap. However, testing for differences BCD showed slightly better predictability than IRM ($p = 0.023$) while the two methods did not differ in reproducibility ($p = 0.15$). On average IRM estimated 29.6 (std=0.83) clusters while BCD estimated 34.8 (0.88) and IDM estimated 17.7 (1.13).

3.3. Varying link density

Figure 6 shows the mean data and its standard error for the reproducibility, predictability, and number of clusters within and between the three datasets when varying link density. The first row shows the Copenhagen dataset, second row Leipzig, third row Beijing, fourth row between dataset. Inspecting clustering reproducibility the general tendency was that BCD per-

28

formed better or on par with IRM across all datasets and for all link densities. For IRM and BCD the general tendency was that reproducibility increased with higher link densities and seems to plateau between 16% and 32%. IDM showed better reproducibility for the lowest link density investigated 2% and then generally decreased with higher link densities and showed better or on par reproducibility compared with IRM and BCD for 2% and 4% link densities.

BCD and IRM showed higher predictability than IDM for all datasets and link densities. For the three within datasets BCD performed better than IRM, for higher link densities these two methods were on par. Please note that the test log likelihood ratio cannot be compared directly between different link densities. BCD generally estimated 5-10 more clusters than IRM. For IRM and BCD the number of clusters estimated increased with increasing link density. The opposite is seen with IDM which estimated fewer clusters with higher link densities.
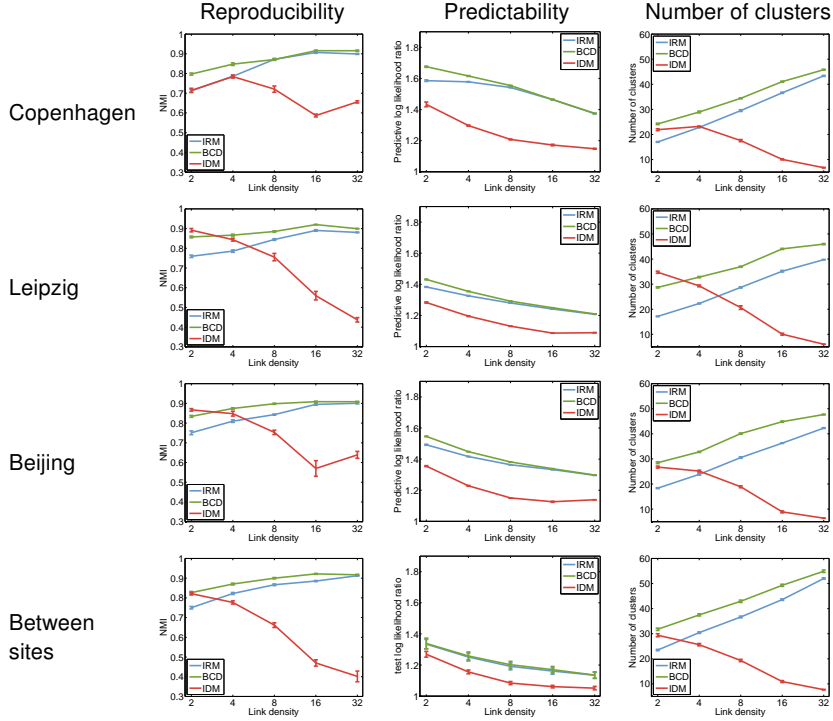
Figure 6: First column normalized mutual information (NMI), second column test log likelihood ratio, and third column number of clusters as function of graph densities. First row Copenhagen, second row Leipzig, third row Beijing, fourth row between sites.

4. Discussion and conclusion

Our aim was to explore statistical models for finding structure in networks at the intermediate level. Accumulated evidence points to the importance of community structure in brain networks, hence, we tested three statistical link models, which differed in terms of the different restrictions that were imposed

30

on how nodes are clustered. The IRM is a very flexible representation for graph clustering, in which nodes can be grouped together without having a high link density among them. The BCD is a constrained version of the IRM that discards such group structures by insisting on higher within-cluster interaction, conforming with the notion of community structure. Finally, the IDM model is further constrained to ignore potential differences in the way nodes in a community interact with other communities, inspired by the methods aimed at identifying structure based on the global modularity concept.

In general IRM and BCD clustered few nodes together corresponding to interhemipheric homologues areas. IRM and BCD model the between-cluster link probabilities, which allows one to inspect how different clusters link to each other, an example of this is shown in figure 4. While a low number of nodes in a specific cluster might not reveal a lot of information in itself important characteristics can be extracted and interpreted when considering the information available from the between-cluster link probabilities. In contrast to these two most expressive models, IDM does not model specific between-cluster link probabilities. This results in larger clusters with relatively high within-cluster link probability, which are grouped together since the model does not care about specific relations to other clusters. These clusters are generally coarser and less nuanced compared to IRM and BCD rendering cluster interpretation difficult. An example of this is the large turquoise cluster shown in figure 3, which is composed of nodes in frontal, occipital, parietal, and temporal lobes.

IRM and BCD produced representations that are predictive and repro-

ducible. Among IRM and BCD, the latter shows better or on par predictabil-ity and reproducibility for all three datasets and between the datasets for all investigated link densities. The predictability of IRM and BCD was clearly better than IDM, which was seen for all datasets and link densities. This suggests that taking between-cluster link probabilities into account improves both reproducibility and data predictability. The clusterings produced by IRM and BCD were very similar with mean NMI between clusterings of 0.94 at 8% link density. The similarity between the representations of IRM and BCD indicates that the flexibility of IRM is not needed when mod-eling rs-fMRI data. Even though IRM is able to cluster nodes such that the clustering does not obey the community structure we see that IRM in general does produce clusterings which are very similar to BCD. The differ-ence between BCD and IRM was most pronounced for smaller link densities suggesting that despite the large similarity between IRM and BCD it helps having the community structure constraint on the clustering. This is most evident for smaller link densities where the graphs contain less information about the network. The better performance of BCD adds to the evidence that coordinated activation in the resting state is community structured.

By invoking a non-parametric Bayesian approach, the three modeling schemes considered are less sensitive to conventional model specification is-sues such as determining the number of communities as the number of clusters is inferred during model inference. However, the models' hyper-parameters still need to be set. In this study uniform priors were used, however other strategies could be considered. For instance, given the Bayesian framework it would be straightforward to sample the hyper-parameters as part of the

model inference (Kemp et al., 2006). Our analysis scheme is a population level model, as we enforced graphs (subjects) to share the model's link probability matrix. This allows us to estimate the out-of-sample predictability for new subjects. We note that it is feasible to model individual link probability matrices (Mørup et al., 2010; Andersen et al., 2012a). This enables test for group differences in link probabilities or correlating with behavioral or personality measures, where specific between-cluster linking can be considered and enables a more specific conclusion about how, e.g., different population groups differ in linking structure.

A number of studies have reported relevance of the conventional network modularity measure to important cognitive measures, such as short term memory capacity, reaction time etc. (Bassett et al., 2011; Stevens et al., 2012; Meunier et al., 2009). Our findings suggest that there is important structure in resting state networks beyond the global modularity. The rich link structures of the relational models can be seen as a way of infering functional integration at the inter-community level as discussed in (Hagmann et al., 2008; Sporns, 2013). Hence, an interesting open question is how to convert the flexible representations of the BCD to summary statistics that can be used as bio-markers. Indeed, initial evidence for the relevance of the community level link density (ρ) as a bio-marker for multiple sclerosis was presented in Mørup et al. (2010)

In conclusion, we evaluated three different Bayesian models for finding structure in rs-fMRI graphs. We showed that BCD performs best compared to IRM and IDM in terms of predictability and reproducibility. This suggests that (1) rs-fMRI data adhere to the community structure and (2) modeling

specific between-cluster linking improves predictability and reproducibility.

Toolbox

A toolbox for performing the experiments conducted in this paper can be found at `https://brainconnectivity.compute.dtu.dk/`

Acknowledgement

5. References

References

Aldous, D., 1985. Exchangeability and related topics. École d'Été de Probabilités de Saint-Flour XIII1983, Lecture Notes in Mathematics 1117, 1–198.

Alexander-Bloch, A.F., Gogtay, N., Meunier, D., Birn, R., Clasen, L., Lalonde, F., Lenroot, R., Giedd, J., Bullmore, E.T., 2010. Disrupted modularity and local connectivity of brain functional networks in childhood-onset schizophrenia. Frontiers in systems neuroscience 4.

Andersen, K.W., Madsen, K.H., Siebner, H., Hansen, L.K., Mørup, M., 2012a. Identification of Functional Clusters in the Striatum Using Infinite Relational Modeling, in: Langs, G., Rish, I., Grosse-Wentrup, M., Murphy, B. (Eds.), Machine Learning and Interpretation in Neuroimaging. Springer Berlin Heidelberg. Lecture Notes in Computer Science, pp. 226–233. doi:`10.1007/978-3-642-34713-9_29`.

Andersen, K.W., Mørup, M., Siebner, H., Madsen, K.H., Hansen, L.K., 2012b. Identifying modular relations in complex brain networks, in: Machine Learning for Signal Processing (MLSP), 2012 IEEE International Workshop on, pp. 1–6. doi:`10.1109/MLSP.2012.6349739`.

Ashburner, J., Friston, K.J., 2005. Unified segmentation. Neuroimage 26, 839–851. doi:`10.1016/j.neuroimage.2005.02.018`.

Barabási, A.L., 2003. Linked: how everything is connected to everything else and what it means for business, science, and everyday life. Plume Editors.

Bassett, D.S., Wymbs, N.F., Porter, M.a., Mucha, P.J., Carlson, J.M., Grafton, S.T., 2011. Dynamic reconfiguration of human brain networks during learning. Proceedings of the National Academy of Sciences of the United States of America 108, 7641–6. doi:`10.1073/pnas.1018985108`.

Birn, R.M., Diamond, J.B., Smith, M.a., Bandettini, P.a., 2006. Separating respiratory-variation-related fluctuations from neuronal-activity-related fluctuations in fMRI. NeuroImage 31, 1536–48.

Biswal, B.B., Mennes, M., Zuo, X.N., Gohel, S., Kelly, C., Smith, S.M., Beckmann, C.F., Adelstein, J.S., Buckner, R.L., Colcombe, S., Dogonowski, A.M., Ernst, M., Fair, D., Hampson, M., Hoptman, M.J., Hyde, J.S., Kiviniemi, V.J., Kötter, R., Li, S.J., Lin, C.P., Lowe, M.J., Mackay, C., Madden, D.J., Madsen, K.H., Margulies, D.S., Mayberg, H.S., McMahon, K., Monk, C.S., Mostofsky, S.H., Nagel, B.J., Pekar, J.J., Peltier, S.J., Petersen, S.E., Riedl, V., Rombouts, S.a.R.B., Rypma, B., Schlaggar, B.L., Schmidt, S., Seidler, R.D., Siegle, G.J., Sorg, C., Teng, G.J., Veijola, J., Villringer, A., Walter, M., Wang, L., Weng, X.C., Whitfield-Gabrieli, S., Williamson, P., Windischberger, C., Zang, Y.F., Zhang, H.Y., Castellanos, F.X., Milham, M.P., 2010. Toward discovery science of human brain function. Proceedings of the National Academy of Sciences of the United States of America 107, 4734–9.

Bullmore, E.T., Bassett, D.S., 2011. Brain graphs: graphical models of the human brain connectome. Annual review of clinical psychology 7, 113–40. doi:`10.1146/annurev-clinpsy-040510-143934`.

Dagli, M.S., Ingeholm, J.E., Haxby, J.V., 1999. Localization of cardiac-induced signal change in fMRI. Neuroimage 9, 407–15.

Dahl, D.B., 2005. Sequentially-Allocated Merge-Split Sampler for Conjugate and Nonconjugate Dirichlet Process Mixture Models. Technical Report, Department of Statistics, Texas A&M University .

Damoiseaux, J.S., Rombouts, S., Barkhof, F., Scheltens, P., Stam, C.J., Smith, S.M., Beckmann, C.F., 2006. Consistent resting-state networks across healthy subjects. Proceedings of the National Academy of Sciences of the United States of America 103, 13848–13853.

Fortunato, S., Barthélemy, M., 2007. Resolution limit in community detection. Proceedings of the National Academy of Sciences of the United States of America 104, 36–41. doi:`10.1073/pnas.0605965104`.

Fox, M., Snyder, A., Vincent, J., 2005. The human brain is intrinsically organized into dynamic, anticorrelated functional networks. Proceedings of the National Academy of Sciences 102, 9673–9678.

Friston, K.J., Williams, S., Howard, R., Frackowiak, R.S., Turner, R., 1996. Movement-related effects in fMRI time-series. Magnetic resonance in medicine 35, 346–55.

Glover, G.H., Li, T.Q., Ress, D., 2000. Image-based method for retrospective correction of physiological motion effects in fMRI: RETROICOR. Magnetic Resonance in Medicine 44, 162–167.

Goutte, C., Toft, P., Rostrup, E., Nielsen, F., Hansen, L.K., 1999. On clustering fMRI time series. NeuroImage 9, 298–310.

Hagmann, P., Cammoun, L., Gigandet, X., Meuli, R., Honey, C.J., Wedeen, V.J., Sporns, O., 2008. Mapping the structural core of human cerebral cortex. PLoS biology 6, e159. doi:`10.1371/journal.pbio.0060159`.

van den Heuvel, M.P., Stam, C.J., Boersma, M., Hulshoff Pol, H.E., 2008. Small-world and scale-free organization of voxel-based resting-state functional connectivity in the human brain. NeuroImage 43, 528–39. doi:`10.1016/j.neuroimage.2008.08.010`.

Jain, S., Neal, R.M., 2004. A Split-Merge Markov chain Monte Carlo Procedure for the Dirichlet Process Mixture Model. Journal of Computational and Graphical Statistics 13, 158–182. doi:`10.1198/1061860043001`.

Kemp, C., Tenenbaum, J., Griffiths, T., Yamada, T., Ueda, N., 2006. Learning systems of concepts with an infinite relational model, in: Proceedings of the National Conference on Artificial Intelligence, Menlo Park, CA; Cambridge, MA; London; AAAI Press; MIT Press; 1999. pp. 381–388.

Lehmann, S., Hansen, L.K., 2007. Deterministic Modularity Optimization. The European Physical Journal B 60, 83–88.

Lund, T.E., Madsen, K.H., Sidaros, K., Luo, W.L., Nichols, T.E., 2006. Non-white noise in fMRI: does modelling have an impact? NeuroImage 29, 54–66. doi:`10.1016/j.neuroimage.2005.07.005`.

McKeown, M., 2003. Independent component analysis of functional MRI: what is signal and what is noise? Current Opinion in Neurobiology 13, 620–629. doi:`10.1016/j.conb.2003.09.012`.

McKeown, M., Makeig, S., Brown, G., 1998. Analysis of fMRI data by blind separation into independent spatial components. Human brain mapping 6, 160–188.

Meunier, D., Achard, S., Morcom, A., Bullmore, E., 2009. Age-related changes in modular organization of human brain functional networks. NeuroImage 44, 715–23. doi:`10.1016/j.neuroimage.2008.09.062`.

Mørup, M., Madsen, K., Dogonowski, A., Siebner, H., Hansen, L., 2010. Infinite relational modeling of functional connectivity in resting state fmri. Advances in Neural Information Processing Systems 23 , 1750–1758.

Mørup, M., Schmidt, M.N., 2012. Bayesian community detection. Neural computation 24, 2434–56. doi:`10.1162/NECO_a_00314`.

Newman, M.E.J., 2006. Modularity and community structure in networks. Proceedings of the National Academy of Sciences of the United States of America 103, 8577–82. doi:`10.1073/pnas.0601602103`.

Nowicki, K., Snijders, T., 2001. Estimation and prediction for stochastic blockstructures. Journal of the American Statistical Association 96, 1077–1087.

Oldfield, R.C., 1971. The assessment and analysis of handedness: The Edinburgh inventory. Neuropsychologia 9, 97–113.

Schmidt, M.N., Mørup, M., 2013. Nonparametric Bayesian Modeling of Complex Networks. IEEE Signal Processing Magazine 30, 110–128.

Smith, A.M., Lewis, B.K., Ruttimann, U.E., Ye, F.Q., Sinnwell, T.M., Yang, Y., Duyn, J.H., Frank, J.A., 1999. Investigation of low frequency drift in fMRI signal. Neuroimage 9, 526–33.

Sporns, O., 2011. The human connectome: a complex network. Annals of the New York Academy of Sciences 1224, 109–25. doi:`10.1111/j.1749-6632.2010.05888.x`.

Sporns, O., 2013. Network attributes for segregation and integration in the human brain. Current Opinion in Neurobiology 23, 162–171. doi:`10.1016/j.conb.2012.11.015`.

Stevens, A.a., Tappon, S.C., Garg, A., Fair, D.a., 2012. Functional brain network modularity captures inter- and intra-individual variation in working memory capacity. PloS one 7, e30468. doi:`10.1371/journal.pone.0030468`.

Strother, S.C., Anderson, J., Hansen, L.K., Kjems, U., Kustra, R., Sidtis, J., Frutiger, S., Muley, S., LaConte, S., Rottenberg, D., 2002. The quantitative evaluation of functional neuroimaging experiments: the NPAIRS data analysis framework. NeuroImage 15, 747–71. doi:`10.1006/nimg.2001.1034`.

Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., Mazoyer, B., Joliot, M., 2002. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. Neuroimage 15, 273–89. doi:`10.1006/nimg.2001.0978`.

Xu, Z., Tresp, V., Yu, K., Kriegel, H., 2006. Infinite hidden relational models, in: In Proceedings of the 22nd International Conference on Uncertainity in Artificial Intelligence (UAI), Citeseer.

Zalesky, A., Fornito, A., Bullmore, E., 2012. On the use of correlation as a measure of network connectivity. Neuroimage 60, 2096–2106. URL: `http://www.sciencedirect.com/science/article/pii/S1053811912001784`, doi:`10.1016/j.neuroimage.2012.02.001`.

Appendix A. FCON1000 data

Data included from Beijing Zang (data included in Beijing_Zang_part2.tar): sub20127, sub20246, sub20765, sub20948, sub21115, sub22201, sub22595, sub22661, sub22715, sub22890, sub26713, sub28206, sub28403, sub28698, sub28792, sub28801, sub28907, sub28965, sub29590, sub29785, sub30272, sub30310, sub30556, sub30616, sub30988, sub31058, sub31729, sub32517, sub32587, sub33747, sub33943, sub33991, sub34895, sub34943, sub35309, sub35776, sub35806, sub36580, sub36942, sub37602, sub38602, sub39725.

Data included from Leipzig: sub00321, sub01002, sub02075, sub07097, sub07374, sub07516, sub07786, sub18698, sub23427, sub25344, sub31577, sub31637, sub36858, sub37308, sub41241, sub47452, sub49383, sub52507, sub52858, sub53063, sub53394, sub59494, sub59709, sub59861, sub61373, sub63957, sub64446, sub72508, sub75022, sub75886, sub77802, sub80206, sub80552, sub85213, sub90843, sub92903, sub94784.

Appendix B. Inference

Appendix B.1. IRM

As stated in section 2.4 the generative model for the Infinite Relational

Model is

Infinite Relational Model	
Cluster assignments:	$\mathbf{z} \sim \mathrm{CRP}(\alpha)$
Link probabilities:	$\rho_{k,l} \sim \mathrm{Beta}(\beta, \beta)$
Links:	$A_{i,j}^{(n)} \sim \mathrm{Bernoulli}(\rho_{z_i,z_j})$

For brevity we define the joint set of graphs as $\mathbf{A} = \{\mathbf{A}^{(1)}, ..., \mathbf{A}^{(N)}\}$. The

Bernoulli likelihood can then be written as:

$$
\begin{aligned}
P(\mathbf{A}|\mathbf{z}, \boldsymbol{\rho}) &= \prod_{n}\prod_{j>i} \rho_{z_i,z_j}^{A_{i,j}^{(n)}}(1 - \rho_{z_i,z_j})^{\left(1 - A_{i,j}^{(n)}\right)} \\
&= \prod_{j>i} \rho_{z_i,z_j}^{\left(\sum_n A_{i,j}^{(n)}\right)}(1 - \rho_{z_i,z_j})^{\left(N - \sum_n A_{i,j}^{(n)}\right)} \\
&= \prod_{k\geq l} \rho_{k,l}^{N_{k,l}^+}(1 - \rho_{k,l})^{N_{k,l}^-},
\end{aligned}
$$

where $N_{k,l}^+$ and $N_{k,l}^-$ is the total number of links and nonlinks for all graphs

between cluster k and l, respectively and N is the number of graphs (sub-

jects). The prior for the link probabilities is a symmetric Beta distribution

and can be written as

$$
P(\boldsymbol{\rho}|\beta) = \prod_{k\geq l} \frac{\Gamma(2\beta)}{\Gamma(\beta)^2} \rho_{k,l}^{\beta-1}(1 - \rho_{k,l})^{\beta-1}
$$

where $\Gamma(x) = (x-1)!$ is the gamma function. The CRP prior for the node partition can be written as

$$P(\mathbf{z}|\alpha) \;=\; \frac{\alpha^K \Gamma(\alpha) \prod_k \Gamma(n_k)}{\Gamma(J+\alpha)}, \tag{B.1}$$

where J is the number of nodes per graph, n_k is the number of nodes assigned to cluster k and K is the number of clusters. These distributions are combined to yield the joint distribution for the IRM:

$$
\begin{aligned}
P(\mathbf{A}, \mathbf{z}, \boldsymbol{\rho}|\alpha, \beta) \;=\;& P(\mathbf{A}|\mathbf{z}, \boldsymbol{\rho}) P(\boldsymbol{\rho}|\beta) P(\mathbf{z}|\alpha) \\
=\;& \Big[\prod_{k \geq l} \rho_{k,l}^{N_{k,l}^+}(1-\rho_{k,l})^{N_{k,l}^-} \Big] \Big[\prod_{k \geq l} \frac{\Gamma(2\beta)}{\Gamma(\beta)^2} \rho_{k,l}^{\beta-1}(1-\rho_{k,l})^{\beta-1} \Big] \Big[\frac{\alpha^K \Gamma(\alpha) \prod_k \Gamma(n_k)}{\Gamma(J+\alpha)} \Big] \\
=\;& \Big[\prod_{k \geq l} \frac{\Gamma(2\beta)}{\Gamma(\beta)^2} \rho_{k,l}^{N_{k,l}^++\beta-1}(1-\rho_{k,l})^{N_{k,l}^-+\beta-1} \Big] \Big[\frac{\alpha^K \Gamma(\alpha) \prod_k \Gamma(n_k)}{\Gamma(J+\alpha)} \Big]
\end{aligned}
$$

Now we can marginalize $\boldsymbol{\rho}$:

$$
\begin{aligned}
P(\mathbf{A}, \mathbf{z}, |\alpha, \beta) \;=\;& \int P(\mathbf{A}, \mathbf{z}, \boldsymbol{\rho}|\alpha, \beta)\mathrm{d}\boldsymbol{\rho} \\
=\;& \Big[\prod_{k \geq l} \frac{\mathrm{B}(N_{k,l}^+ + \beta, N_{k,l}^- + \beta)}{\mathrm{B}(\beta, \beta)} \Big] \Big[\frac{\alpha^K \Gamma(\alpha) \prod_k \Gamma(n_k)}{\Gamma(J+\alpha)} \Big]
\end{aligned}
$$

where $\mathrm{B}(x, y) = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}$ is the Beta function. Finally using Bayes' theorem we can find the posterior distribution of the assignment of a single node z_i

$$P(z_i = l|\mathbf{A}, \mathbf{z}_{\backslash i}, \beta, \alpha) = \frac{P(\mathbf{A}, \mathbf{z}_{\backslash i}, z_i = l|\alpha, \beta)}{\sum_{l'} P(\mathbf{A}, \mathbf{z}_{\backslash i}, z_i = l'|\alpha, \beta)}$$

where $\mathbf{z}_{\backslash i}$ is the assignments of all nodes except node i. By writing out this equation and finding parts which change when a node is assigned to a cluster (Schmidt and Mørup, 2013) we have that:

$$P(z_i = l|\mathbf{A}, \mathbf{z}_{\backslash i}, \beta, \alpha) \;\propto\; \begin{cases} n_{l \backslash i} \prod_k \frac{\mathrm{B}(N_{k,l}^{+\backslash i}+r_{i,k}^+\beta, N_{k,l}^{-\backslash i}+r_{i,k}^-+\beta)}{\mathrm{B}(N_{k,l}^{+\backslash i}+\beta, N_{k,l}^{-\backslash i}+\beta)} & \text{if } n_{l\backslash i} > 0 \\ \alpha \prod_k \frac{\mathrm{B}(r_{i,k}^+\beta, r_{i,k}^-+\beta)}{\mathrm{B}(\beta, \beta)} & \text{otherwise.} \end{cases} \tag{B.2}$$

$N_{k,l}^{+\backslash i}$ and $N_{k,l}^{-\backslash i}$ is the number of links and nonlinks between clusters k and l not counting links from node i. $n_{l\backslash i}$ is the number of nodes assigned to cluster l disregarding the assignment of node i. $r_{i,k}^+$ and $r_{i,k}^-$ is the number of links and nonlinks from node i to any node in cluster k. This posterior likelihood can be evaluated efficiently since we only need to compute $\mathbf{N}^+$ and $\mathbf{N}^-$ and evaluate the Beta function for entries affected by the considered assignment change. The posterior likelihood is used in the Gibbs sampler to infer the node assignments.

Appendix B.2. IDM

The generative model for the Infinite Diagonal Model is given by:

Infinite Diagonal Model

$$\text{Cluster assignments:} \quad \mathbf{z} \sim \text{CRP}(\alpha)$$

$$\text{Link probabilities:} \quad \rho_{k,l} \sim \begin{cases} \rho_k = \text{Beta}(\beta, \beta) & \text{if } k = l \\ \rho_b = \text{Beta}(\beta, \beta) & \text{otherwise.} \end{cases}$$

$$\text{Links:} \quad A_{i,j}^{(n)} \sim \text{Bernoulli}(\rho_{z_i, z_j})$$

The Bernoulli likelihood can accordingly be written as:

$$P(\mathbf{A}|\mathbf{z}, \boldsymbol{\rho}) = \rho_b^{N_b^+} (1 - \rho_b)^{N_b^-} \Big[\prod_k \rho_k^{N_k^+} (1 - \rho_k)^{N_k^-} \Big],$$

where N_k^+ and N_k^- is the number of links and nonlinks within cluster k and N_b^+ and N_b^- is the total number of links and nonlinks which fall between clusters. The prior for the link probabilities can be written as

$$P(\boldsymbol{\rho}|\beta) = \frac{\Gamma(2\beta)}{\Gamma(\beta)^2} \rho_b^{\beta-1}(1 - \rho_b)^{\beta-1} \Big[\prod_k \frac{\Gamma(2\beta)}{\Gamma(\beta)^2} \rho_k^{\beta-1}(1 - \rho_k)^{\beta-1} \Big]$$

44

809 The prior for the node partition is the same as the IRM model (equation
810 B.1). The joint distribution for the IDM can then be written as:

$$
\begin{aligned}
811 \quad P(\mathbf{A}, \mathbf{z}, \boldsymbol{\rho} | \alpha, \beta) &= P(\mathbf{A}|\mathbf{z}, \boldsymbol{\rho}) P(\boldsymbol{\rho}|\beta) P(\mathbf{z}|\alpha) \\
812 &= \rho_b^{N_b^+} (1 - \rho_b)^{N_b^-} \Big[\prod_k \rho_k^{N_k^+} (1 - \rho_k)^{N_k^-} \Big] \\
813 &\quad \frac{\Gamma(2\beta)}{\Gamma(\beta)^2} \rho_b^{\beta-1} (1 - \rho_b)^{\beta-1} \Big[\prod_k \frac{\Gamma(2\beta)}{\Gamma(\beta)^2} \rho_k^{\beta-1} (1 - \rho_k)^{\beta-1} \Big] \Big[\frac{\alpha^K \Gamma(\alpha) \prod_k \Gamma(n_k)}{\Gamma(J + \alpha)} \Big] \\
814 &= \frac{\Gamma(2\beta)}{\Gamma(\beta)^2} \rho_b^{N_b^+ + \beta - 1} (1 - \rho_b)^{N_b^- + \beta - 1} \\
815 &\quad \Big[\prod_k \frac{\Gamma(2\beta)}{\Gamma(\beta)^2} \rho_k^{N_k^+ + \beta - 1} (1 - \rho_k)^{N_k^- + \beta - 1} \Big] \Big[\frac{\alpha^K \Gamma(\alpha) \prod_k \Gamma(n_k)}{\Gamma(J + \alpha)} \Big]
\end{aligned}
$$

816 Now marginalizing over $\boldsymbol{\rho}$:

$$
\begin{aligned}
817 \quad P(\mathbf{A}, \mathbf{z}, |\alpha, \beta) &= \int P(\mathbf{A}, \mathbf{z}, \boldsymbol{\rho} | \alpha, \beta) \mathrm{d}\boldsymbol{\rho} \\
818 &= \frac{\mathrm{B}(N_b^+ + \beta, N_b^- + \beta)}{\mathrm{B}(\beta, \beta)} \Big[\prod_k \frac{\mathrm{B}(N_k^+ + \beta, N_k^- + \beta)}{\mathrm{B}(\beta, \beta)} \Big] \Big[\frac{\alpha^K \Gamma(\alpha) \prod_k \Gamma(n_k)}{\Gamma(J + \alpha)} \Big]
\end{aligned}
$$

819 Finally using Bayes' theorem we can find the posterior distribution of the
820 assignment of a single node z_i

$$
821 \quad P(z_i = l | \mathbf{A}, \mathbf{z}_{\backslash i}, \beta, \alpha) = \frac{P(\mathbf{A}, \mathbf{z}_{\backslash i}, z_i = l | \alpha, \beta)}{\sum_{l'} P(\mathbf{A}, \mathbf{z}_{\backslash i}, z_i = l' | \alpha, \beta)}
$$

822 By writing out this equation and finding parts which change when a node is
823 assigned to a cluster we find that:

$$
824 \quad P(z_i = l | \mathbf{A}, \mathbf{z}_{\backslash i}, \beta, \alpha) \propto
$$

$$
825 \quad \begin{cases} n_{l \backslash i} \dfrac{\mathrm{B}(N_l^{+\backslash i} + r_{i,l}^+ + \beta, N_l^{-\backslash i} + r_{i,l}^- + \beta)}{\mathrm{B}(N_l^{+\backslash i} + \beta, N_l^{-\backslash i} + \beta)} \dfrac{\mathrm{B}(N_b^{+\backslash i} + \sum_{k \neq l} r_{i,k}^+ + \beta, N_b^{-\backslash i} + \sum_{k \neq l} r_{i,k}^- + \beta)}{\mathrm{B}(N_b^{+\backslash i} + \beta, N_b^{-\backslash i} + \beta)} & \text{if } n_{l \backslash i} > 0 \\ \alpha \dfrac{\mathrm{B}(N_b^{+\backslash i} + \sum_{k \neq l} r_{i,k}^+ + \beta, N_b^{-\backslash i} + \sum_{k \neq l} r_{i,k}^- + \beta)}{\mathrm{B}(N_b^{+\backslash i} + \beta, N_b^{-\backslash i} + \beta)} & \text{otherwise.} \end{cases} \tag{B.3}
$$

826 $r_{i,l}^+$ and $r_{i,l}^-$ is the number of links and nonlinks from node i to any node in
827 cluster l.

45

Appendix B.3. BCD

This section will give a short description of the inference in the Bayesian

Community Detection (BCD) model, for further details please refer to Mørup

and Schmidt (2012). The generative model for BCD is given by:

Bayesian Community Detection

$$\text{Cluster assignments}: \quad \mathbf{z} \sim \text{CRP}(\alpha)$$

$$\text{Cluster gap}: \quad \gamma \sim \text{Beta}(v, v)$$

$$\text{Link probability}: \quad \rho_{k,l} \sim \begin{cases} \text{Beta}(\beta, \beta) & \text{if } k = l \\ \text{BetaInc}(\beta, \beta, w_{k,l}) & \text{otherwise.} \end{cases}$$

$$\text{where } w_{k,l} = \min[\gamma\rho_{ll}, \gamma\rho_{kk}]$$

$$\text{Links}: \quad A_{i,j}^{(n)} \sim \text{Bernoulli}(\rho_{z_i, z_j})$$

If we let $\dot{\boldsymbol{\rho}} = \{\rho_{k,l} | k = l\}$ and $\ddot{\boldsymbol{\rho}} = \{\rho_{k,l} | k \neq l\}$ be the set of within and

between link probabilities respectively. Then the joint distribution can be

written as

$$
\begin{aligned}
P(\mathbf{A}, \mathbf{z}, \boldsymbol{\rho}, \gamma | \alpha, \beta) &= P(\mathbf{A}|\mathbf{z}, \boldsymbol{\rho})P(\ddot{\boldsymbol{\rho}}|\dot{\boldsymbol{\rho}}, \gamma, \beta)P(\dot{\boldsymbol{\rho}}|\beta)P(\gamma|v)P(\mathbf{z}|\alpha) \\
&= \Big[\prod_{n=1}^{N} \prod_{j>i} \rho_{z_i, z_j}^{A_{i,j}^{(n)}} (1 - \rho_{z_i, z_j})^{1 - A_{i,j}^{(n)}} \Big] \\
&\quad \times \Big[\prod_{k>l} \frac{\rho_{k,l}^{\beta-1}(1 - \rho_{k,l})^{\beta-1}}{\text{B}_{x_{k,l}}(\beta, \beta)} \Big] \Big[\prod_{l=1}^{K} \frac{\rho_{l,l}^{\beta-1}(1 - \rho_{l,l})^{\beta-1}}{\text{B}(\beta, \beta)} \Big] \\
&\quad \times \Big[\frac{\gamma^{v-1}(1 - \gamma)^{v-1}}{\text{B}(v, v)} \Big] \Big[\frac{\alpha^K \Gamma(\alpha) \prod_k \Gamma(n_k)}{\Gamma(J + \alpha)} \Big]
\end{aligned}
$$

Integrating over $\ddot{\boldsymbol{\rho}}$:

$$P(\mathbf{A}, \mathbf{z}, \dot{\boldsymbol{\rho}}, \gamma | \alpha, \beta) = \int P(\mathbf{A}, \mathbf{z}, \boldsymbol{\rho}, \gamma | \alpha, \beta) \mathrm{d}\ddot{\boldsymbol{\rho}}$$

$$= \Big[\prod_{k=1}^{K} \frac{\rho_{k,k}^{N_{k,k}^{+}+\beta-1} (1 - \rho_{k,k})^{N_{k,k}^{-}+\beta-1}}{\mathrm{B}(\beta, \beta)} \Big]$$

$$\times \Big[\prod_{k>l} \frac{\mathrm{B}_{x_{k,l}}(N_{k,l}^{+} + \beta, N_{k,l}^{-} + \beta)}{\mathrm{B}_{x_{k,l}}(\beta, \beta)} \Big]$$

$$\times \Big[\frac{\gamma^{v-1}(1 - \gamma)^{v-1}}{\mathrm{B}(v, v)} \Big] \Big[\frac{\alpha^{K} \Gamma(\alpha) \prod_{k} \Gamma(n_{k})}{\Gamma(J + \alpha)} \Big]$$

Again, using Bayes theorem and eliminating terms which does not depend on $\rho_{l,l}$ the marginal posterior reduces to

$$P(\rho_{l,l} | \mathbf{A}, \mathbf{z}, \dot{\boldsymbol{\rho}} \backslash \rho_{l,l}, \beta, \alpha, \gamma) \propto \rho_{l,l}^{N_{l,l}^{+}+\beta-1} (1 - \rho_{l,l})^{N_{l,l}^{-}+\beta-1} \prod_{k \neq l} \frac{\mathrm{B}_{x_{l,k}}(N_{k,l}^{+}+\beta, N_{k,l}^{-}+\beta)}{\mathrm{B}_{x_{k,l}}(\beta, \beta)}$$

The conditional distribution for a node assignment is given as (Mørup and Schmidt, 2012):

$$P(z_i = l | \mathbf{A}, \mathbf{z}_{\backslash i}, \dot{\boldsymbol{\rho}}, \beta, \alpha, \gamma) \propto \rho_{l,l}^{r_{i,l}^{+}} (1 - \rho_{l,l})^{r_{i,l}^{-}} \alpha^{K} n_{l \backslash i}$$

$$\prod_{k \neq l} \frac{\mathrm{B}_{x_{k,l}}(N_{k,l}^{+\backslash i} + r_{i,k}^{+} + \beta, N_{k,l}^{-\backslash i} + r_{i,k}^{-} + \beta)}{\mathrm{B}_{x_{k,l}}(N_{k,l}^{+\backslash i} + \beta, N_{k,l}^{-\backslash i} + \beta)}$$

When terms which does not depend on γ are ignored the posterior reduces to

$$P(\gamma | \mathbf{A}, \mathbf{z}, \dot{\boldsymbol{\rho}}, \beta, \alpha) \propto \gamma^{v-1}(1 - \gamma)^{v-1} \prod_{k>l} \frac{\mathrm{B}_{x_{k,l}}(N_{k,l}^{+\backslash i} + \beta, N_{k,l}^{-\backslash i} + \beta)}{\mathrm{B}_{x_{k,l}}(\beta, \beta)}$$

Appendix C. Clusters labels

47

IRM	BCD	IDM
Frontal Mid Orb L	Frontal Mid Orb L	Frontal Mid Orb L
Frontal Mid Orb R	Frontal Mid Orb R	Frontal Mid Orb R
Frontal Sup Orb L	Frontal Sup Orb L	Frontal Sup Orb L
Frontal Sup Orb R	Frontal Sup Orb R	Frontal Sup Orb R
Hippocampus L	Hippocampus L	Hippocampus L
Hippocampus R	Hippocampus R	Hippocampus R
ParaHippocampal L	ParaHippocampal L	ParaHippocampal L
ParaHippocampal R	ParaHippocampal R	ParaHippocampal R
Thalamus L	Thalamus L	Thalamus L
Thalamus R	Thalamus R	Thalamus R
Parietal Inf L	Parietal Inf L	Parietal Inf L
Parietal Inf R	Parietal Inf R	Parietal Inf R
Temporal Pole Sup L	Temporal Pole Sup L	Temporal Pole Sup L
Temporal Pole Sup R	Temporal Pole Sup R	Temporal Pole Sup R
Calcarine L	Calcarine L	Calcarine L
Calcarine R	Calcarine R	Calcarine R
Lingual L	Lingual L	Lingual L
Lingual R	Lingual R	Lingual R
Cingulum Mid L	Cingulum Mid L	Cingulum Mid L
Cingulum Mid R	Cingulum Mid R	Cingulum Mid R
Cuneus L	Cuneus L	Cuneus L
Cuneus R	Cuneus R	Cuneus R
Occipital Sup L	Occipital Sup L	Occipital Sup L
Occipital Sup R	Occipital Sup R	Occipital Sup R
Occipital Inf L	Occipital Inf L	Occipital Inf L
Occipital Inf R	Occipital Inf R	Occipital Inf R
Occipital Mid L	Occipital Mid L	Occipital Mid L
Occipital Mid R	Occipital Mid R	Occipital Mid R
Fusiform L	Fusiform L	Fusiform L
Fusiform R	Fusiform R	Fusiform R
Paracentral Lobule L	Paracentral Lobule L	Paracentral Lobule L
Paracentral Lobule R	Paracentral Lobule R	Paracentral Lobule R
Postcentral L	Postcentral L	Postcentral L
Postcentral R	Postcentral R	Postcentral R
Precentral L	Precentral L	Precentral L
Precentral R	Precentral R	Precentral R
Supp Motor Area L	Supp Motor Area L	Supp Motor Area L
Supp Motor Area R	Supp Motor Area R	Supp Motor Area R
Parietal Sup L	Parietal Sup L	Parietal Sup L
Parietal Sup R	Parietal Sup R	Parietal Sup R
Precuneus L	Precuneus L	Precuneus L
Precuneus R	Precuneus R	Precuneus R
Temporal Inf L	Temporal Inf L	Temporal Inf L
Temporal Inf R	Temporal Inf R	Temporal Inf R
Temporal Sup L	Temporal Sup L	Temporal Sup L
Temporal Sup R	Temporal Sup R	Temporal Sup R
Temporal Mid L	Temporal Mid L	Temporal Mid L
Temporal Mid R	Temporal Mid R	Temporal Mid R
Frontal Mid L	Frontal Mid L	Frontal Mid L
Frontal Mid R	Frontal Mid R	Frontal Mid R
Cingulum Ant L	Cingulum Ant L	Cingulum Ant L
Cingulum Ant R	Cingulum Ant R	Cingulum Ant R
Frontal Inf Oper L	Frontal Inf Oper L	Frontal Inf Oper L
Frontal Inf Oper R	Frontal Inf Oper R	Frontal Inf Oper R
Frontal Inf Tri L	Frontal Inf Tri L	Frontal Inf Tri L
Frontal Inf Tri R	Frontal Inf Tri R	Frontal Inf Tri R
Frontal Sup L	Frontal Sup L	Frontal Sup L
Frontal Sup R	Frontal Sup R	Frontal Sup R
Frontal Sup Medial L	Frontal Sup Medial L	Frontal Sup Medial L
Frontal Sup Medial R	Frontal Sup Medial R	Frontal Sup Medial R
Frontal Inf Orb L	Frontal Inf Orb L	Frontal Inf Orb L
Frontal Inf Orb R	Frontal Inf Orb R	Frontal Inf Orb R

IRM	BCD	IDM
Insula L	Insula L	Insula L
Insula R	Insula R	Insula R
Rolandic Oper L	Rolandic Oper L	Rolandic Oper L
Rolandic Oper R	Rolandic Oper R	Rolandic Oper R
Heschl L	Heschl L	Heschl L
Heschl R	Heschl R	Heschl R
SupraMarginal L	SupraMarginal L	SupraMarginal L
SupraMarginal R	SupraMarginal R	SupraMarginal R
Caudate L	Caudate L	Caudate L
Caudate R	Caudate R	Caudate R
Putamen L	Putamen L	Putamen L
Putamen R	Putamen R	Putamen R
Angular L	Angular L	Angular L
Angular R	Angular R	Angular R
Cingulum Post L	Cingulum Post L	Cingulum Post L
Cingulum Post R	Cingulum Post R	Cingulum Post R
Frontal Med Orb L	Frontal Med Orb L	Frontal Med Orb L
Frontal Med Orb R	Frontal Med Orb R	Frontal Med Orb R
Rectus L	Rectus L	Rectus L
Rectus R	Rectus R	Rectus R
Cerebelum 7b L	Cerebelum 7b L	Vermis 9
Cerebelum 7b R	Cerebelum 7b R	Cerebelum 7b R
Vermis 7	Cerebelum 10 L	Cerebelum 10 L
Vermis 9	Cerebelum 10 R	Cerebelum 10 R
Cerebelum 9 L	Cerebelum 3 L	Cerebelum 3 L
Cerebelum 9 R	Cerebelum 3 R	Vermis 3
Vermis 3	Vermis 3	Vermis 3
Amygdala L	Amygdala L	Amygdala L
Amygdala R	Amygdala R	Amygdala R
Temporal Pole Mid L	Temporal Pole Mid L	Temporal Pole Mid L
Temporal Pole Mid R	Temporal Pole Mid R	Temporal Pole Mid R
Olfactory L	Olfactory L	Olfactory L
Olfactory R	Olfactory R	Olfactory R
Pallidum L	Pallidum L	Pallidum L
Pallidum R	Pallidum R	Pallidum R
Vermis 10	Vermis 10	Vermis 10
Vermis 1 2	Vermis 1 2	Vermis 1 2
Cerebelum 3 R	Cerebelum 9 L	Cerebelum 9 L
Cerebelum 10 L	Cerebelum 9 R	Cerebelum 9 R
Cerebelum 10 R	Vermis 7	Cerebelum 3 R
Vermis 8	Vermis 8	Vermis 8
Vermis 3	Vermis 6	Cerebelum 7b L
Cerebelum 4 5 L	Cerebelum 4 5 L	Cerebelum 4 5 L
Cerebelum 4 5 R	Cerebelum 4 5 R	Cerebelum 4 5 R
Vermis 4 5	Vermis 4 5	Vermis 4 5
Vermis 6	Vermis 6	Vermis 6
Cerebelum 8 L	Cerebelum 8 L	Cerebelum 8 L
Cerebelum 8 R	Cerebelum 8 R	Cerebelum 8 R
Cerebelum 6 L	Cerebelum 6 L	Cerebelum 6 L
Cerebelum 6 R	Cerebelum 6 R	Cerebelum 6 R
Cerebelum Crus1 L	Cerebelum Crus1 L	Cerebelum Crus1 L
Cerebelum Crus1 R	Cerebelum Crus1 R	Cerebelum Crus1 R
Cerebelum Crus2 L	Cerebelum Crus2 L	Cerebelum Crus2 L
Cerebelum Crus2 R	Cerebelum Crus2 R	Cerebelum Crus2 R

Figure C.7: Labels from the extracted clusters using IRM, BCD, and IDM. The colors correspond to the clusters from figure 3.

Joint Modelling of Structural and Functional Brain Networks

Andersen, K. W., Herlau, T., Mørup, M., Schmidt, M. N., Madsen, K. H., Dyrby, T. D., Lyksborg, M., Siebner, H. R., Hansen, L. K., (2014) 'Joint Modelling of Structural and Functional Brain Networks', in preparation.

Joint Modelling of Structural and Functional Brain Networks

Kasper Winther Andersen[1,2,3], Tue Herlau[1,3], Morten Mørup[1], Mikkel N. Schmidt[1], Kristoffer H. Madsen[2], Tim B. Dyrby[2], Mark Lyksborg[1,2], Hartwig Siebner[2], and Lars Kai Hansen[1]

[1] DTU Compute, Technical University of Denmark
[2] Danish Research Centre for Magnetic Resonance, Copenhagen University Hospital Hvidovre
[3] shared first author

Abstract. Functional and structural magnetic resonance imaging (fMRI and dMRI) have become the most important noninvasive windows into the human brain. A major challenge in the analysis of brain networks is to establish the similarities and dissimilarities between functional and structural connectivity. We formulate a non-parametric Bayesian network model which allows for joint modelling and integration of multiple network types. We demonstrate the model's ability to detect vertices that share structure across networks, both on artificial data as well as joint analysis of fMRI and dMRI data. Using two fMRI and dMRI scans per subject, we found that structure in dMRI networks are very consistent across scans, while fMRI graphs only show intermediate consistency. Comparing fMRI and dMRI we found limited overlap in shared structure.

Keywords: Functional magnetic resonance imaging, diffusion weighted imaging, Bayesian network models, complex networks

1 Introduction

While the dominant paradigm in neuroimaging remains functional localization, there is much current interest in understanding the mechanisms behind brain wide coordination in more complex human behaviors. Network representations and graph theoretical analyses offer new means to understand functional coordination both under normal behaviors and pathology. Current brain imaging technology offers multiple views on these networks and it has been suggested that structural information obtained via diffusion weighted imaging and information from functional MRI (fMRI) can be combined in a synergistic way to produce a more complete picture, see e.g., [1–7].

We note that there are major systematic differences to the networks derived from structural and functional imaging methods. While functional connectivity obtained by fMRI measures dependencies on a time scale of 1-100 seconds, diffusion imaging reveals 'wiring' that supports direct communication between neurons at time scales on the order of 1-100 milliseconds. Furthermore, the characterization of connectivity by fMRI is confounded by the presence of vascular fluctuations, movement and other confounding signals. Diffusion imaging based tractography is challenged in several ways,

both in the detection of local connectivity due to presence of crossing/kissing fibers and in accurate detection of long-range connections due to long path lengths and the presence of funneling structures such as the corpus callosum [8] making tracking of in particular interhemispheric connections difficult.

Existing work on combining functional connectivity and structural connectivity is based on comparing or correlating results obtained from each method independently [9, 10]. This includes comparisons of descriptive measures of structural and functional networks like the distribution of motifs or other global properties, such as the link distribution, small world properties, or degree of modularity, for a recent review see [7]. However, in order to fully benefit from the individual advantages of each modality thereby improving detection power and explicitly capture shared network structure, models that are able to integrate both types on information in the same framework are needed. This is indeed the aim of recent work on informing functional clustering by measures of structural connectivity as derived from diffusion MRI [11]. Their so-called anatomically weighted functional clustering method (awFC) produces more strongly autocorrelated functional networks compared with conventional unweighted functional clustering of time series.

The above mentioned approaches for structure-function integration assume that the relation is *global*, leading, e.g., [11] to an awFC model in which the integration is controlled by a single tunable integration strength parameter for the whole brain volume, not allowing for partial integration and non-stationarity across brain structures. Such assumed homogeneity is convenient from a computational point of view but is can hardly be justified in our knowledge about distributed brain function or structural properties. In general anatomy and function are markedly inhomogeneous: Areas differ much in physiology, the type of signal they handle, and their role. Thus, there is significant differentiation across the volume leading us to hypothesize a structure-function relation which is non-stationary.

In order to address this issue we set up an expressive generative model that can potentially learn the structure-function relation locally, and we compare this flexible model with two reference models based on no coupling and global coupling respectively. The new model is formulated in a new Bayesian non-parametric multi-graph modeling framework that allows for arbitrary relations between a set of modules also determined by data and most importantly directly models the potential integration between structure and function at the node level. In particular we propose a non-parametric generative model that divides nodes into two types *i) shared nodes that exhibit the same clustering structure across modalities* and *ii) individual nodes (not shared) that are clustered differently across modalities*. This admit the analysis of shared and individual structures in networks with multiple link types, i.e., graphs defined by functional and structural connectivity respectively (see also Figure 1). This admits the investigation of the structure-function relationship in brain networks and provides a general framework for multimodal integration.

To validate our approach we train and test on two separate sets of structural and functional MRI graphs obtained for a total of 22 subjects. We compare our modeling to a model that assumes there are no clusters shared between the functional and structural connectivity graphs and a model that assumes clusters are fully shared between the
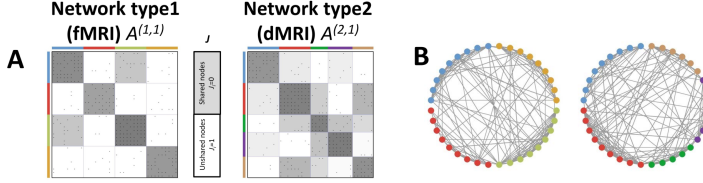
Fig. 1. Schematic model description. Panel A shows the adjacency matrices for two network types. The small black dots indicate edges. Grouping structures are indicated with colors to the left and above the adjacency matrices and group link probabilities are indicated with the gray-scaled background. In this example nodes belonging to the first two modules are shared across network type while the remaining nodes show different grouping in the two networks. Panel B gives a graphical layout of the networks where nodes are shown with filled circles and edges as lines. The color indicate the same grouping structure as in panel A.

functional and structural connectivity graphs. In addition, we investigate if subjects are better modelled by subject specific partitions of the brain regions compared to partitions that are the same across subjects.

1.1 Non-parametric multigraph models

A functional and a structural connectivity graph defined on the same set of R brain regions can be jointly represented as an edge-labelled multigraph with two edge labels or, equivalently, as an $(R \times R \times 2)$ adjacency tensor, $A_{i,j,k}$. The infinite relational model (IRM) [12, 13] is a prominent non-parametric Bayesian modeling approach for analyzing such multigraphs, extending its parametric counterpart, the stochastic block-model [14–16]. These models describe a graph as blocks of vertices which exhibit similar patterns of relations to vertices in other blocks.

For a single unipartite graph with adjacency matrix $A_{i,j}$, the IRM model is defined by the following generative process: The vertices are partitioned according to a Chinese restaurant process (CRP), $z \sim \mathrm{CRP}(\alpha)$; for each pair of clusters an edge probability is generated from a Beta distribution, $\eta_{\mu,\nu} \sim \mathrm{Beta}(\gamma, \delta)$; and each possible edge is then generated from a Bernoulli distribution, $A_{i,j} \sim \mathrm{Bernoulli}(\eta_{z_i,z_j})$.

Kemp et al. [12] describe IRM models applied to clustering multiple graphs, multiple types of relations, and hypergraphs where edges can connect more than two vertices. Based on the IRM model, Mørup et al. [17] model the functional connectivity of the brain in resting state across multiple subjects in a model where the clustering is shared between subjects but edge probabilities are allowed to vary between subjects. Ishiguro et al. [18] extend the IRM model to the analysis of time-evolving networks: Here the edge probabilities are assumed to be stationary whereas the clustering structure is modelled as time-dependent, inspired by the nonparametric hidden Markov model [19]. Based on an Indian buffet process and a logistic linear model, Miller et al. [20] model multiple graphs in a flexible framework where each vertex is allowed to belong to multiple clusters.

4

---

1) Draw the probability of nodes being shared:
$$\phi \sim \mathrm{Beta}(\kappa, \ell)$$
2) For each node draw a binary variable J_i indicating whether the node is shared or individual:
$$\forall i \; J_i \sim \mathrm{Bernoulli}(\phi)$$
3) For the shared nodes draw cluster assignments:
$$\forall i : J_i = 0 \; \{z_i^{(0)}\} \sim \mathrm{CRP}(\alpha)$$
4) For the individual nodes draw cluster assignments:
$$\forall n, i : J_i = 1 \; \{z_i^{(n)}\} \sim \mathrm{CRP}(\alpha)$$
5) For individual and shared clusters draw link probability:
$$\forall n, m, \mu, \nu : \mu \le \nu \; \eta_{\mu,\nu}^{(n,m)} \sim \mathrm{Beta}(\gamma, \delta)$$
6) For all subjects and network types draw links:
$$\forall n, m, i, j : i < j \; A_{i,j}^{(n,m)} \sim \mathrm{Bernoulli}(\eta_{z_i^{(n \cdot J_i)}, z_j^{(n \cdot J_j)}}^{(n,m)})$$

Table 1. The proposed generative model

While all these existing models are able to detect modular structure in multiple networks, none of them address the question of determining whether or not common patterns of relations exist in a set of graphs (see Figure 1). In this work we wish to jointly model graphs derived from structural and functional brain imaging with the purpose of identifying which parts of the graphs that exhibit similar patterns and which parts that are distinct in the two modalities. We present an approach to modeling multiple graphs in which patterns explicitly are either shared or individual for each type of graph.

2 Methods

Consider a set of graphs $\{G_{n,m}(V, E_{n,m}) : n \in \{1, \ldots, N\}, m \in \{1, \ldots, M\}\}$, each defined on the same set of vertices $V = \{v_1, \ldots, v_K\}$. The set of graphs is indexed by (n, m) where the former in the present context refers to whether the graph is derived from functional or diffusion MRI ($N = 2$) and the latter refers to different subjects or repetitions.

2.1 Generative model

We propose the following model for the set of graphs: For each vertex v_i, a Bernoulli variable J_i indicates whether its linking properties are shared across graphs, $J_i = 0$, or are individual for each graph, $J_i = 1$. According to a Chinese restaurant process (CRP), the set of shared vertices $V_{\mathrm{sh.}} = \{v_i : J_i = 0\}$ is divided into a partition Z_0 represented by $\{z_i^{(0)}\}_{i \in V_{\mathrm{sh.}}}$ where $z_i^{(0)}$ is an index of the cluster to which vertex v_i belongs. For each $n \in \{1, \ldots, N\}$ the set of individual nodes $V_{\mathrm{in.}} = \{v_i : J_i = 1\}$ is divided into a partition Z_n represented by $\{z_i^{(n)}\}_{i \in V_{\mathrm{in.}}}$. For each graph in the set, a link probability variable $\eta_{\mu,\nu}^{(n,m)}$ determines the probability of observing a link $A_{i,j}^{(n,m)}$ in the graph $G_{n,m}$ between two vertices v_i and v_j belonging to cluster μ and ν respectively.

The full generative model is described in Table 1. The tunable parameters of the model are $\psi = \{\kappa, \ell, \alpha, \gamma, \delta\}$ and they were all set to unity in our experiments.

The model posits a stochastic block structure which is identical across subjects / repetitions (m) but only partially shared across network types (n). Since the degree of shared block structure is variable, the model allows us to learn from data to what extent common block structure is present across different types of networks. We further consider two special cases of this partially-shared block model: A fully-shared $J_i = 1 \, \forall i$ and a fully-unshared $J_i = 0 \, \forall i$ model. In the fully-shared model a single partitioning of the nodes is shared between all graphs; however, the link probabilities remain separate for each graph allowing the model to capture shared group structure despite differences in link densities. This type of modeling has previously been considered in [20]. In the fully-unshared model each graph is assumed independent, and our model reduces to N independent multi-graph infinite relational models (IRM) [12].

2.2 Inference

The parameters ϕ and $\eta_{\mu,\nu}^{(n,m)}$ are marginalized analytically, resulting in the following joint posterior distribution, known up to a multiplicative constant,

$$
p(z, J | A, \psi) \propto \left[\prod_{(m,n)} \prod_{(\mu,\nu)} \frac{\mathrm{B}\left(\gamma + E_{\mu,\nu}^{(m,n)}, \delta + \overline{E}_{\mu,\nu}^{(m,n)}\right)}{\mathrm{B}(\gamma, \delta)} \right]
$$
$$
\times \left[\frac{\Gamma\left(|V_{\mathrm{in.}}| + \alpha\right)}{\Gamma\left(|V_{\mathrm{sh.}}| + \alpha\right)} \prod_{n=0}^{N} \frac{\alpha^{|Z_n|} \Gamma\left(|Z_n|\right)}{\Gamma\left(|V_{\mathrm{in.}}| + \alpha\right)} \prod_{\zeta \in Z_n} \Gamma\left(|\zeta|\right) \right] \times \frac{\mathrm{B}\left(\kappa + |V_{\mathrm{in.}}|, \ell + |V_{\mathrm{sh.}}|\right)}{\mathrm{B}(\kappa, \ell)},
$$

where $\mathrm{B}(\cdot)$ is the Beta function, ζ is a set of vertices belonging to a given cluster, and $E_{\mu,\nu}^{(m,n)}$ ($\overline{E}_{\mu,\nu}^{(m,n)}$) denotes the number of edges (non-edges) between pairs of vertices in cluster μ and ν respectively within the graph $G_{n,m}$. Evaluating this expression entails counting these numbers of edges and non-edges as well as the number of vertices associated with each cluster.

Since computing the entire posterior distribution of the remaining parameters in the model is not computationally tractable (the number of possible values of the discrete parameters in z and J is huge even for small graphs), we resort to inference by Markov chain Monte Carlo methods. In particular, we use joint Gibbs sampling of $\left(J_i, z_i^{(0)}, \ldots, z_i^{(N)}\right)$. For each update we thus need to consider $|Z_0| + 1 + \prod_{n=1}^{N} \left(|Z_n| + 1\right)$ possible cluster assignments (where $|Z_n|$ is computed excluding v_i). For large N the $z_i^{(n)}$ could be Gibbs sampled individually reducing the number of possible assignments to $|Z_0| + |Z_n| + 2$. In addition to this we use two types of split-merge sampling inspired by the approach of Jain and Neal [21], where splitting or merging clusters is used as a Metropolis-Hastings proposal. In the first type, we consider splitting or merging clusters within a single partition Z_n. In the second type, we consider merging all vertices in a non-shared cluster Z_n into a shared cluster in Z_0 or vice versa splitting a shared cluster into N individual clusters. Inference in the two special cases, the fully-shared and fully-unshared models, follows trivially.

To decrease burnin time we used a tempered distribution in the burning phase proportional to $\exp(\tau_t \log p(z, J|A, \psi))$ where τ is the temperature schedule. In our experiments we choose $\tau_t = \tau_0 + (1 - \tau_0)(t/T)^{3/2}$ as the temperature schedule with $\tau_0 = 0.05/|J|$ and T is the total number of burnin iterations.

2.3 Model validation

To validate our models on functional and structural connectivity graphs we use split-half cross-validation. For each subject we generate two sets of functional and structural connectivity graphs. We infer parameters on one set of fMRI and dMRI graphs (i.e., training data) and evaluate the inferred models on the other set of fMRI and dMRI graphs (i.e., test data). We will use two metrics to assess performance; area under the receiver operating characteristic curve (AUC) and mutual information (MI). AUC will be used to assess the model's ability to account for structure in the test fMRI and dMRI graphs whereas MI will be used to evaluate how reproducible the clustering structure is. We used $\mathrm{MI}(z^{(1)}, z^{(2)}) = \sum_{\mu\mu'} p(\mu, \mu') \log \frac{p(\mu, \mu')}{p(\mu)p(\mu')}$. Both AUC and MI will be calculated for the fMRI and dMRI test data separately.

2.4 Data acquisition

We use fMRI and dMRI data obtained from 22 healthy volunteers.

Functional MRI: Resting state functional magnetic resonance imaging (rs-fMRI) data was recorded for 20 min (482 volumes) per subject. The first two volumes were discarded to account for T1 equilibrium effects, the remaining 480 volumes were realigned to the time-series mean and spatially normalized to the MNI template using SPM8. Nuisance effects related to residual movement or physiological effects were removed using a linear filter comprised of 24 motion related and a total of 64 physiological effects including cardiac, respiratory and respiration volume over time as well as the time series from CSF and white matter voxels in both the left and right hemispheres. In order to enable the characterization of reproducibility and predictive likelihood, the fMRI data was blocked into 6 equally sized blocks each consisting of 80 non-overlapping volumes. For each subject the 6 blocks was randomly split into two independent datasets each containing 3 blocks. For each of the two datasets the mean signal in each of the 116 regions as defined in the AAL database [22] was extracted and a $[116 \times 116]$ correlation matrix was formed for each of the datasets for each subject.

Diffusion MRI: Each subject underwent two diffusion weighted imaging (DWI) sessions. For each session diffusion along 61 directions were recorded with $b = 1200s/mm^2$. Additionally, 10 $b = 0$ images were obtained. To compensate for subject motion and Eddy currents artifacts, the DWI volumes of each subject were aligned with the first DWI volume (b-value=0) using an affine image transformation with the cost criteria of normalized mutual information (NMI) [23]. The displacements of the affine transformations were combined with the displacement maps of SPM8's Fieldmap approach [24], and displacement maps correcting for the non-linearity of the scanner gradients,

resulting in a single resampling for each volume, achieved using cubic interpolation. The rotational parts of the affine transformations are extracted and gradient directions corrected using the Finite Strain approach [25]. FSL's Bedpostx was used to estimate voxelwise diffusion parameters for each subjects' two DWI sessions. Bedpostx uses Markov Chain Monte Carlo sampling to build distributions of the diffusion parameters and allows for detection of crossing fibres. FSL's probtrackx [26] was used for probabilistic tractography. The AAL brain regions were normalized to each subject's native space and used for seeding and target regions. For each voxel a distribution of 5000 pathway samples was generated and the average count of pathways to each region was used to generate the connectivity matrix C. In order to generate un-directional graphs we average $\tilde{C} = \frac{1}{2}(C + C^\top)$.

Construction of graphs: The correlation matrix derived from functional MRI and the averaged pathway count matrix derived from diffusion MRI were thresholded to form binary adjacency matrices. The threshold was selected such that the resulting graph had a specific link density: We experimented with graph densities between 3 % and 50 %.

3 Results and discussion

3.1 Synthetic data analysis

To evaluate the method, we considered networks with an underlying structure supportive of a partially shared node configuration, but with a varying difficulty parameterized by an order parameter ϵ, and each of the 3 models under consideration, the fully-shared, fully-unshared and partially-shared, was evaluated on the networks.

For each ϵ a network of K vertices and C communities was constructed as follows: $\frac{n}{2}$ elements of J was assigned to 1. The two relations $z^{(1)}$ and $z^{(2)}$ was constructed as C clusters of equal size and a community assignment consistent with J according to the generative model. The $\eta^{(n)}$-matrices were constructed as a linear combination $\eta^{(n)} = at + b(1 - t)$ where t is the $C \times C$ identity matrix, and a, b was selected under the constraints (i) conditional on $\eta^{(n)}$, the networks should all have a constant average density and (ii) if $\epsilon = 0$ all entries of $\eta^{(n)}$ should be equal and if $\epsilon = 1$ then $b = 0$, see inserts in figure 2 for 3 networks corresponding to ϵ values indicated by black vertical lines. We selected $C = 8, K = 128$ and an average density of 0.1. As a measure of performance we considered the Normalized Mutual Information (NMI) (NMI$(z', z'') = \frac{2\text{MI}(z',z'')}{\text{MI}(z',z')+\text{MI}(z'',z'')}$) between inferred $z^{(n)}$-matrices and their true value, and for the partially-shared model also NMI between J and its true value.

Results can be seen in figure 2. As expected the fully-unshared and partially-shared model both recover the true community structure for $\epsilon \approx 1$, while the fully-shared model cannot properly represent the true structure in this limit and hence detect poorer cluster structure. However for very hard problems ($\epsilon \approx 0.25$), only the fully-shared communities present sufficient evidence to infer the cluster structure, and thus the shared model will *begin* to detect earlier than the fully-unshared model.

The partially-shared model is also able to exploit this feature of the data and hence begin to detect earlier. Interestingly the partially-shared model seem to outperform the
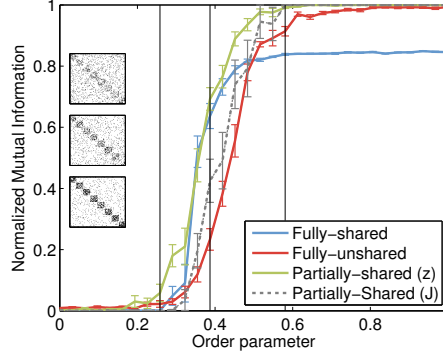
Fig. 2. Result of simulation of artificial data. Each value is averaged over 10 networks and each simulation is run for 60 burn-in and 10 actual samples. All models were initialized in a fully independent state with random cluster assignments across 8 clusters.

fully-shared model even for low ϵ, this may be due to each correctly identified clusters making the remaining search space smaller, hence also enabling detection on the part of the structure which is not shared.

3.2 Analysis of fMRI and dMRI data

We consider analysis of the MR data where fMRI and dMRI are modelled jointly for each subject individually (i.e. $M = 1$). In addition we consider an within-modality model, which models fMRI1 vs. fMRI2 and dMRI1 vs. dMRI2. For the latter analysis we are not able to calculate AUC nor MI since we use both scans in the same model.

The results of the analysis is given in figure 3 where the mean and standard error across subjects of the fraction of shared vertices, AUC, number of clusters, MI, and NMI are given. The figure reveals a high consistency in the dMRI graphs between scans with a mean fraction of shared nodes above 90% for all link densities. fMRI shows less consistency in clustering across the two scans. The mean fraction of shared nodes is 90% at 3% link density, which falls below 50% for 50% link density. The second row plots the AUC for training the model on set1 and predicting on test2 (left) and vica virca (right). We observe that for link densities $\geq 10\%$ the partially-shared model is on par with the fully-unshared model. The third row plots the number clusters found by the model, which increases with increasing link density and is slightly higher for fMRI than dMRI. In the fourth row we plot the MI and NMI between clusterings in the two scans. We find in general higher clustering consistency in dMRI than fMRI, which is also consistent with the finding that dMRI in general have a higher number of shared nodes in the within-modality model. The fully-unshared model have higher NMI for

the dMRI data and is on par with the partially-shared for fMRI. The plots for AUC and number of clusters show great consistency across the two scans.

4   Conclusion

Integration of the different modalities of structural and functional connectivity data remain a difficult challenge in computational neuroscience. We have proposed a framework which allows for joint modelling of different network types and inference of vertices which share structure. The model was evaluated on artificial data demonstrating its ability to correctly identify shared and unshared vertices as well as indicating an improved detection threshold in a high-noise setting compared to individual modelling.

The model was evaluated on fMRI and dMRI data derived from 22 subjects. By using two scans of each data type per subject we ran the proposed model both within- and between-modality. In the within-modality we observed that the dMRI exhibit a highly consistent structure across data sets. The number of shared nodes were for all link densities greater than 90%. The fMRI graphs were less consistent and the number of shared nodes decreased when increasing link density. When comparing fMRI and dMRI we found a low degree of similarity in clustering as revealed by the low number of shared nodes. For link densities $\geq 10\%$ the fraction of shared nodes were below 25%. There were a higher between-modality consistency for low (3% and 5%) link densities this, however, it most likely due large null-clusters, which is large clusters with very low link densities both within and between clusters.

Our model provides a principled unsupervised framework for establishing structure-function relationship in brain networks. We believe this is a crucial property for models which purpose to investigate the relationship between these data modalities and advance our understanding on the relationship between the functional and structural organization of the brain in both healthy and unhealthy subject populations.

Here we used the model to find similarity between fMRI and dMRI graphs. However, the model is general and can be used for any types of graphs defined on the same set of nodes. Thus, the model could be used to infer group differences in clustering and thus potentially serve as bio-marker for psychiatric and neurological disorders.

References

1. Bullmore, E., Sporns, O.: Complex brain networks: graph theoretical analysis of structural and functional systems. Nature Reviews Neuroscience **10**(3) (2009) 186–198
2. Guye, M., Bettus, G., Bartolomei, F., Cozzone, P.J.: Graph theoretical analysis of structural and functional connectivity MRI in normal and pathological brain networks. Magma (New York, N.Y.) (March 2010)
3. Zhang, Z., Liao, W., Chen, H., Mantini, D., Ding, J., Xu, Q., Wang, Z., Yuan, C., Chen, G., Jiao, Q., et al.: Altered functional–structural coupling of large-scale brain networks in idiopathic generalized epilepsy. Brain **134**(10) (2011) 2912–2928
4. Johansen-Berg, H., Behrens, T.E.J., Robson, M.D., Drobnjak, I., Rushworth, M.F.S., Brady, J.M., Smith, S.M., Higham, D.J., Matthews, P.M.: Changes in connectivity profiles define functionally distinct regions in human medial frontal cortex. Proceedings of the National Academy of Sciences of the United States of America **101**(36) (2004) 13335

5. Honey, C.J., Sporns, O., Cammoun, L., Gigandet, X., Thiran, J.P., Meuli, R., Hagmann, P.: Predicting human resting-state functional connectivity from structural connectivity. Proceedings of the National Academy of Sciences **106**(6) (2009) 2035

6. Knösche, T.R., Tittgemeyer, M.: The role of long-range connectivity for the characterization of the functional–anatomical organization of the cortex. Frontiers in systems neuroscience **5** (2011)

7. Sporns, O.: The human connectome: a complex network. Annals of the New York Academy of Sciences **1224**(1) (2011) 109–125

8. Häberling, I.S., Badzakova-Trajkov, G., Corballis, M.C.: Callosal tracts and patterns of hemispheric dominance: A combined fmri and dti study. Neuroimage **54**(2) (2011) 779–786

9. Lowe, M.J., Beall, E.B., Sakaie, K.E., Koenig, K.A., Stone, L., Marrie, R.A., Phillips, M.D.: Resting state sensorimotor functional connectivity in multiple sclerosis inversely correlates with transcallosal motor pathway transverse diffusivity. Hum.Brain Mapp. **29**(1097-0193 (Electronic) LA - eng PT - Journal Article PT - Research Support, Non-U.S. Gov't SB - IM) (July 2008) 818–827

10. Filippi, M., Agosta, F.: Structural and functional network connectivity breakdown in Alzheimers disease studied with magnetic resonance imaging techniques. Journal of Alzheimer's disease : JAD **24**(3) (January 2011) 455–74

11. Bowman, F.D., Zhang, L., Derado, G., Chen, S.: Determining Functional Connectivity usying fMRI Data with Diffusion-Based Anatomical Weighting. NeuroImage (May 2012)

12. Kemp, C., Tenenbaum, J.B., Griffiths, T.L., Yamada, T., Ueda, N.: Learning systems of concepts with an infinite relational model. In: Proceedings of the national conference on artificial intelligence. Volume 21. (2006) 381

13. Xu, Z., Tresp, V., Yu, K., Kriegel, H.P.: Learning infinite hidden relational models. Uncertainity in Artificial Intelligence (UAI2006) (2006)

14. Holland, P.W., Blackmond, K., Leinhardt, S.: Stochastic blockmodels: First steps. Social networks **5**(2) (1983) 109–137

15. Snijders, T.A.B., Nowicki, K.: Estimation and prediction for stochastic blockmodels for graphs with latent block structure. Journal of Classification **14**(1) (1997) 75–100

16. White, H.C., Boorma, S.A., Breiger, R.L.: Social structure from multiple networks. American Journal of Sociology **81** (1976) 730–780

17. Mørup, M., Madsen, K.H., Dogonowski, A., Siebner, H., Hansen, L.K.: Infinite relational modeling of functional connectivity in resting state fMRI. In: Advances in Neural Information Processing Systems (NIPS). (2010) 1750–1758

18. Ishiguro, K., Iwata, T., Ueda, N., Tenenbaum, J.B.: Dynamic infinite relational model for time-varying relational data analysis. In: Advances in Neural Information Processing Systems (NIPS). (2010) 919–927

19. Teh, Y.W., Jordan, M.I., Beal, M.J., Blei, D.M.: Hierarchical Dirichlet processes. Journal of the American Statistical Association **101**(476) (2006) 1566–1581

20. Miller, K.T., Griffiths, T.L., Jordan, M.I.: Nonparametric latent feature models for link prediction. Advances in Neural Information Processing Systems (NIPS) (2009) 1276–1284

21. Jain, S., Neal, R.M.: A split-merge Markov chain Monte Carlo procedure for the Dirichlet process mixture model. Journal of Computational and Graphical Statistics **13**(1) (2004) 158–182

22. Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., Mazoyer, B., Joliot, M.: Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. NeuroImage **15**(1) (January 2002) 273–289

23. Collignon, A., Maes, F., Delaere, D., Vandermeulen, D., Suetens, P., Marchal, G.: Automated multi-modality image registration based on information theory. In: Information processing in medical imaging. Volume 3. (1995) 264–274

24. Jezzard, P., Balaban, R.S.: Correction for geometric distortion in echo planar images from bo field variations. Magn Reson Med **34**(1) (1995) 65–73
25. Alexander, D.C., Pierpaoli, C., Basser, P.J., Gee, J.C.: Spatial transformations of diffusion tensor magnetic resonance images. Medical Imaging, IEEE Transactions on
26. Behrens, T.E.J., Woolrich, M.W., Jenkinson, M., Johansen-Berg, H., Nunes, R.G., Clare, S., Matthews, P.M., Brady, J.M., Smith, S.M.: Characterization and propagation of uncertainty in diffusion-weighted MR imaging. Magnetic resonance in medicine: Official journal of the Society of Magnetic Resonance in Medicine / Society of Magnetic Resonance in Medicine **50**(5) (November 2003) 1077–88
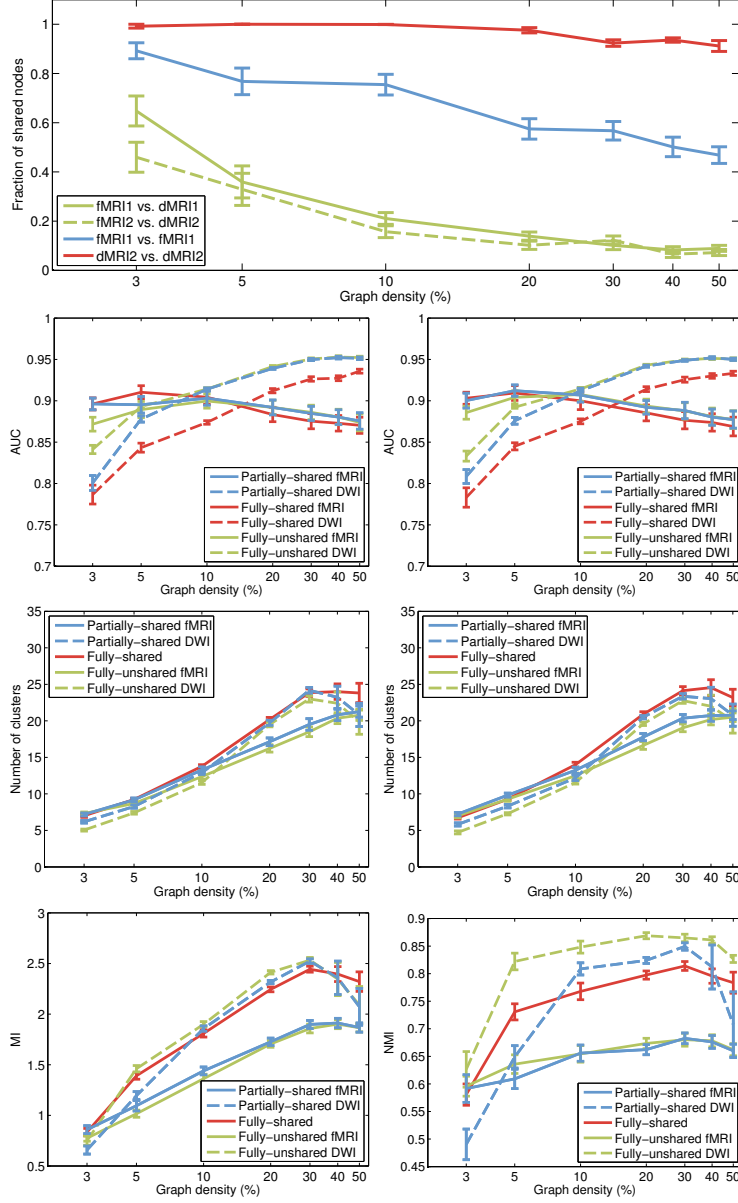
Fig. 3. Model inferred on individual subjects.

Bibliography

Abrahamsen, T. J. and Hansen, L. K. (2011), 'Sparse non-linear denoising: Generalization performance and pattern reproducibility in functional MRI', *Pattern Recognition Letters* **32**(15), 2080–2085.

Aldous, D. J. (1985), 'Exchangeability and related topics', *École d'Été de Probabilités de Saint-Flour XIII - 1983 (Lecture Notes in Mathematics)* **1117**, 1–198.

Alexander-Bloch, A. F., Gogtay, N., Meunier, D., Birn, R., Clasen, L., Lalonde, F., Lenroot, R., Giedd, J. and Bullmore, E. T. (2010), 'Disrupted modularity and local connectivity of brain functional networks in childhood-onset schizophrenia', *Frontiers in systems neuroscience* **4**.

Alexander, D., Pierpaoli, C., Basser, P. and Gee, J. (2001), 'Spatial transformations of diffusion tensor magnetic resonance images', *IEEE transactions on medical imaging* **20**(11), 1131–1139.

Allen, E. A., Damaraju, E., Plis, S. M., Erhardt, E. B., Eichele, T. and Calhoun, V. D. (2012), 'Tracking whole-brain connectivity dynamics in the resting state', *Cerebral Cortex* pp. 1–14.

Andersen, K. W., Madsen, K. H., Siebner, H. R., Schmidt, M. N., Mørup, M. and Hansen, L. K. (2013), 'Coordinated activation in the resting brain is community structured', *submitted* .

Andersen, K. W., Mørup, M., Siebner, H., Madsen, K. H. and Hansen, L. K. (2012), Identifying modular relations in complex brain networks, *in* '2012 IEEE International Workshop on Machine Learning for Signal Processing', IEEE, pp. 1–6.

Ashburner, J. and Friston, K. J. (2005), 'Unified segmentation.', *NeuroImage* **26**(3), 839–51.

Bassett, D. S., Wymbs, N. F., Porter, M. a., Mucha, P. J., Carlson, J. M. and Grafton, S. T. (2011), 'Dynamic reconfiguration of human brain networks during learning.', *Proceedings of the National Academy of Sciences of the United States of America* **108**(18), 7641–6.

Beauducel, A. (2001), 'Problems with parallel analysis in data sets with oblique simple structure', *Methods of Psychological Research Online* **6**(2).

Behrens, T. E. J., Woolrich, M. W., Jenkinson, M., Johansen-Berg, H., Nunes, R. G., Clare, S., Matthews, P. M., Brady, J. M. and Smith, S. M. (2003), 'Characterization and propagation of uncertainty in diffusion-weighted MR imaging.', *Magnetic resonance in medicine : official journal of the Society of Magnetic Resonance in Medicine / Society of Magnetic Resonance in Medicine* **50**(5), 1077–88.

Birn, R. M., Diamond, J. B., Smith, M. A. and Bandettini, P. A. (2006), 'Separating respiratory-variation-related fluctuations from neuronal-activity-related fluctuations in fMRI.', *NeuroImage* **31**(4), 1536–48.

Bishop, C. M. (2006), *Pattern Recognition and Machine Learning (Information Science and Statistics)*, Springer-Verlag New York, Inc., Secaucus, NJ, USA.

Biswal, B., Yetkin, F. Z., Haughton, V. M. and Hyde, J. S. (1995), 'Functional connectivity in the motor cortex of resting human brain using echo-planar MRI', *Magn Reson.Med.* **34**, 537–541.

Bowman, F. D., Zhang, L., Derado, G. and Chen, S. (2012), 'Determining functional connectivity using fMRI data with diffusion-based anatomical weighting.', *NeuroImage* **62**(3), 1769–1779.

Brandes, U., Delling, D., Gaertler, M., Hoefer, M., Nikoloski, Z. and Wagner, D. (2007), 'On Finding Graph Clusterings with Maximum modularity', *Graph-Theoretic Concepts in Computer Science* **4769**, 121–132.

Bullmore, E. T. and Bassett, D. S. (2011), 'Brain graphs: graphical models of the human brain connectome.', *Annual review of clinical psychology* **7**, 113–40.

Collignon, A., Maes, F., Delaere, D., Vandermeulen, D., Suetens, P. and Marchal, G. (1995), 'Automated multi-modality image registration based on information theory', *Information processing in medical imaging* **3**, 264–274.

Compston, A. and Coles, A. (2002), 'Multiple sclerosis.', *Lancet* **359**(9313), 1221–31.

Confavreux, C. and Vukusic, S. (2006), 'Natural history of multiple sclerosis: a unifying concept.', *Brain : a journal of neurology* **129**(Pt 3), 606–16.

Dagli, M. S., Ingeholm, J. E. and Haxby, J. V. (1999), 'Localization of cardiac-induced signal change in fMRI.', *Neuroimage* **9**, 407–15.

Damoiseaux, J. S. and Greicius, M. D. (2009), 'Greater than the sum of its parts: a review of studies combining structural connectivity and resting-state functional connectivity.', *Brain structure & function* **213**(6), 525–33.

Damoiseaux, J. S., Rombouts, S., Barkhof, F., Scheltens, P., Stam, C. J., Smith, S. M. and Beckmann, C. F. (2006), 'Consistent resting-state networks across healthy subjects', *Proceedings of the National Academy of Sciences of the United States of America* **103**(37), 13848–13853.

Di Martino, A., Scheres, A., Margulies, D. S., Kelly, A. M. C., Uddin, L. Q., Shehzad, Z., Biswal, B., Walters, J. R., Castellanos, F. X. and Milham, M. P. (2008), 'Functional connectivity of human striatum: a resting state FMRI study.', *Cerebral cortex (New York, N.Y. : 1991)* **18**(12), 2735–47.

Dogonowski, A.-M. (2012), Resting state Functional Connectivity of the Motor System in Multiple Sclerosis, PhD thesis.

Dogonowski, A.-M., Andersen, K. W., Madsen, K. H., Sørensen, P. S., Paulson, O. B., Blinkenberg, M. and Siebner, H. R. (2014), 'Multiple sclerosis impairs regional functional connectivity in the cerebellum', *NeuroImage: Clinical* **4**, 130–138.

Dogonowski, A.-M., Siebner, H. R., Sørensen, P. S., Paulson, O. B., Dyrby, T. B., Blinkenberg, M. and Madsen, K. H. (2013), 'Resting-state connectivity of pre-motor cortex reflects disability in multiple sclerosis.', *Acta neurologica Scandinavica* **128**, 328–335.

Dogonowski, A.-M., Siebner, H. R., Sørensen, P. S., Wu, X., Biswal, B., Paulson, O. B., Dyrby, T. B., Skimminge, A., Blinkenberg, M. and Madsen, K. H. (2013), 'Expanded functional coupling of subcortical nuclei with the motor resting-state network in multiple sclerosis.', *Multiple sclerosis (Houndmills, Basingstoke, England)* **19**(5), 559–66.

Du, W., Calhoun, V. D., Li, H., Ma, S., Eichele, T., Kiehl, K. a., Pearlson, G. D. and Adali, T. (2012), 'High classification accuracy for schizophrenia with rest and task FMRI data.', *Frontiers in human neuroscience* **6**(June), 145.

Friston, K. J., Williams, S., Howard, R., Frackowiak, R. S. and Turner, R. (1996), 'Movement-related effects in fMRI time-series.', *Magnetic resonance in medicine* **35**(3), 346–55.

Gamboa, O. L., Tagliazucchi, E., von Wegner, F., Jurcoane, A., Wahl, M., Laufs, H. and Ziemann, U. (2013), 'Working memory performance of early MS patients correlates inversely with modularity increases in resting state functional connectivity networks.', *NeuroImage* pp. 1–11.

Glover, G. H., Li, T. Q. and Ress, D. (2000), 'Image-based method for retrospective correction of physiological motion effects in fMRI: RETROICOR', *Magnetic Resonance in Medicine* **44**(1), 162–167.

Goldenberg, M. M. (2012), 'Multiple sclerosis review', *P & T : a peer-reviewed journal for formulary management* **37**(3), 175–84.

Greicius, M. D., Supekar, K., Menon, V. and Dougherty, R. F. (2009), 'Resting-state functional connectivity reflects structural connectivity in the default mode network.', *Cerebral cortex (New York, N.Y. : 1991)* **19**(1), 72–8.

Haber, S. (2003), 'The primate basal ganglia: parallel and integrative networks', *Journal of Chemical Neuroanatomy* **26**(4), 317–330.

He, Y., Wang, L., Zang, Y., Tian, L., Zhang, X., Li, K. and Jiang, T. (2007), 'Regional coherence changes in the early stages of Alzheimer's disease: a combined structural and resting-state functional MRI study.', *NeuroImage* **35**(2), 488–500.

Hinne, M., Ambrogioni, L., Janssen, R. J., Heskes, T. and van Gerven, M. a. J. (2014), 'Structurally-informed Bayesian functional connectivity analysis.', *NeuroImage* **86**, 294–305.

Hofman, J. M. and Wiggins, C. H. (2008), 'A Bayesian approach to network modularity', *Phys. Rev. Lett* **100**(25), 1–4.

Honey, C. J., Sporns, O., Cammoun, L., Gigandet, X., Thiran, J. P., Meuli, R. and Hagmann, P. (2009), 'Predicting human resting-state functional connectivity from structural connectivity.', *Proceedings of the National Academy of Sciences of the United States of America* **106**(6), 2035–40.

Horn, J. (1965), 'A rationale and test for the number of factors in factor analysis', *Psychometrika* **30**(2), 179–185.

Jain, S. and Neal, R. M. (2004), 'A Split-Merge Markov chain Monte Carlo Procedure for the Dirichlet Process Mixture Model', *Journal of Computational and Graphical Statistics* **13**(1), 158–182.

Jezzard, P. and Balaban, R. S. (1995), 'Correction for geometric distortion in echo planar images from B0 field variations.', *Magnetic resonance in medicine* **34**(1), 65–73.

Jørgensen, K. W. and Hansen, L. K. (2012), 'Model Selection for Gaussian Kernel PCA Denoising', *Neural Networks and Learning Systems, IEEE Transactions on* **23**(1), 163–168.

Karrer, B. and Newman, M. (2011), 'Stochastic blockmodels and community structure in networks', *Physical Review E* **83**(1).

Kemp, C., Tenenbaum, J., Griffiths, T., Yamada, T. and Ueda, N. (2006), Learning systems of concepts with an infinite relational model, *in* 'Proceedings of the 21th National Conference on Artificial Intelligence (AAAI)', Menlo Park, CA; Cambridge, MA; London; AAAI Press; MIT Press; 1999, pp. 381–388.

Kendall, M. and Smith, B. (1939), 'The problem of m rankings', *The annals of mathematical statistics* **10**(3), 275–287.

Kurtzke, J. F. (1983), 'Rating Neurologic Impairment in Multiple-Sclerosis - An Expanded Disability Status Scale (EDSS)', *Neurology* **33**(11), 1444–1452.

Ledesma, R. and Valero-Mora, P. (2007), 'Determining the Number of Factors to Retain in EFA: an easy-to-use computer program for carrying out Parallel Analysis', *Practical Assessment, Research and* **12**(2).

Legendre, P. (2005), 'Species associations: the Kendall coefficient of concordance revisited', *Journal of Agricultural, Biological, and Environmental Statistics* **10**(2), 226–245.

Lehmann, S. and Hansen, L. K. (2007), 'Deterministic Modularity Optimization', *The European Physical Journal B* **60**(1), 83–88.

Liang, P., Liu, Y., Jia, X., Duan, Y., Yu, C., Qin, W., Dong, H., Ye, J. and Li, K. (2011), 'Regional homogeneity changes in patients with neuromyelitis optica revealed by resting-state functional MRI.', *Clinical neurophysiology : official journal of the International Federation of Clinical Neurophysiology* **122**(1), 121–7.

Liu, H., Liu, Z., Liang, M., Hao, Y., Tan, L., Kuang, F., Yi, Y., Xu, L. and Jiang, T. (2006), 'Decreased regional homogeneity in schizophrenia: a resting state functional magnetic resonance imaging study.', *Neuroreport* **17**(1), 19–22.

Liu, J. (1996), 'Nonparametric hierarchical Bayes via sequential imputations', *The Annals of Statistics* **24**(3), 911–930.

Lowe, M. J., Phillips, M. D., Lurito, J. T., Mattson, D., Dzemidzic, M. and Mathews, V. P. (2002), 'Multiple sclerosis: low-frequency temporal blood oxygen level-dependent fluctuations indicate reduced functional connectivity initial results.', *Radiology* **224**, 184–92.

Lublin, F. D. and Reingold, S. C. (1996), 'Defining the clinical course of multiple sclerosis : results of an international survey', *Neurology* **46**, 907–911.

Lund, T. E., Madsen, K. H., Sidaros, K., Luo, W.-L. and Nichols, T. E. (2006), 'Non-white noise in fMRI: does modelling have an impact?', *NeuroImage* **29**(1), 54–66.

Meunier, D., Achard, S., Morcom, A. and Bullmore, E. (2009), 'Age-related changes in modular organization of human brain functional networks.', *NeuroImage* **44**(3), 715–23.

Mika, S., Schölkopf, B., Smola, A., Müller, K., Scholz, M. and Rätsch, G. (1999), 'Kernel PCA and de-noising in feature spaces', *Advances in neural information processing systems* **11**(1), 536–542.

Mørup, M., Madsen, K. H., Dogonowski, A.-M., Siebner, H. and Hansen, L. K. (2010), 'Infinite relational modeling of functional connectivity in resting state fMRI', *Advances in Neural Information Processing Systems 23* pp. 1750–1758.

Mørup, M. and Schmidt, M. N. (2012), 'Bayesian community detection.', *Neural computation* **24**(9), 2434–56.

Mørup, M., Schmidt, M. N. and Hansen, L. K. (2011), Infinite multiple membership relational modeling for complex networks, *in* 'Machine Learning for Signal Processing (MLSP), 2011 IEEE International Workshop on'.

Newman, M. E. J. (2006), 'Modularity and community structure in networks.', *Proceedings of the National Academy of Sciences of the United States of America* **103**(23), 8577–82.

Newman, M. and Girvan, M. (2004), 'Finding and evaluating community structure in networks', *Physical Review E* **69**(2), 026113.

Ng, B., Varoquaux, G., Poline, J. B. and Thirion, B. (2013), Implications of Inconsistencies between fMRI and dMRI on Multimodal Connectivity Estimation, *in* 'MICCAI - 16th International Conference on Medical Image Computing and Computer Assisted Intervention - 2013 (2013)', Vol. 2013.

Ng, B., Varoquaux, G., Poline, J. and Thirion, B. (2012), A Novel Sparse Graphical Approach for Multimodal Brain Connectivity Inference, *in* 'Medical Image Computing and Computer-Assisted Intervention–MICCAI 2012', pp. 707–714.

Nowicki, K. and Snijders, T. (2001), 'Estimation and prediction for stochastic blockstructures', *Journal of the American Statistical Association* **96**(455), 1077–1087.

Ogawa, S., Lee, T. M., Kay, a. R. and Tank, D. W. (1990), 'Brain magnetic resonance imaging with contrast dependent on blood oxygenation.', *Proceedings of the National Academy of Sciences of the United States of America* **87**(24), 9868–72.

Oldfield, R. C. (1971), 'The assessment and analysis of handedness: The Edinburgh inventory', *Neuropsychologia* **9**(1), 97–113.

Polman, C. H., Reingold, S. C., Edan, G., Filippi, M., Hartung, H. P., Kappos, L., Lublin, F. D., Metz, L. M., McFarland, H. F., O'Connor, P. W., Sandberg-Wollheim, M., Thompson, A. J., Weinshenker, B. G. and Wolinsky, J. S. (2005), 'Diagnostic criteria for multiple sclerosis: 2005 Revisions to the "McDonald Criteria"', *Annals of Neurology* **58**(6), 840–846.

Rasmussen, P. M., Abrahamsen, T. J., Madsen, K. H. and Hansen, L. K. (2012), 'Nonlinear denoising and analysis of neuroimages with kernel principal component analysis and pre-image estimation.', *NeuroImage* **60**(3), 1807–18.

Reese, T. G., Heid, O., Weisskoff, R. M. and Wedeen, V. J. (2003), 'Reduction of eddy-current-induced distortion in diffusion MRI using a twice-refocused spin echo.', *Magnetic resonance in medicine : official journal of the Society of Magnetic Resonance in Medicine / Society of Magnetic Resonance in Medicine* **49**(1), 177–82.

Roosendaal, S. D., Schoonheim, M. M., Hulst, H. E., Sanz-Arigita, E. J., Smith, S. M., Geurts, J. J. G. and Barkhof, F. (2010), 'Resting state networks change in clinically isolated syndrome.', *Brain : a journal of neurology* **133**(Pt 6), 1612–21.

Sidhu, G. S., Asgarian, N., Greiner, R. and Brown, M. R. G. (2012), 'Kernel Principal Component Analysis for dimensionality reduction in fMRI-based diagnosis of ADHD.', *Frontiers in systems neuroscience* **6**(November), 74.

Skudlarski, P., Jagannathan, K., Anderson, K., Stevens, M. C., Calhoun, V. D., Skudlarska, B. a. and Pearlson, G. (2010), 'Brain connectivity is not only lower but different in schizophrenia: a combined anatomical and functional approach.', *Biological psychiatry* **68**(1), 61–9.

Smith, A. M., Lewis, B. K., Ruttimann, U. E., Ye, F. Q., Sinnwell, T. M., Yang, Y., Duyn, J. H. and Frank, J. A. (1999), 'Investigation of low frequency drift in fMRI signal.', *Neuroimage* **9**, 526–33.

Snyder, A. Z. and Raichle, M. E. (2012), 'A brief history of the resting state: The Washington University perspective.', *NeuroImage* **62**(2), 902–10.

Song, X., Chen, N.-K. and Gaur, P. (2014), 'A kernel machine-based fMRI physiological noise removal method.', *Magnetic resonance imaging* **32**(2), 150–62.

Song, X., Ji, T. and Wyrwicz, A. (2008), Baseline drift and physiological noise removal in high field fmri data using kernel pca, *in* 'Acoustics, Speech and Signal Processing, 2008. ICASSP 2008. IEEE International Conference on.', pp. 441–444.

Sorg, C., Riedl, V., Mühlau, M., Calhoun, V. D., Eichele, T., Läer, L., Drzezga, A., Förstl, H., Kurz, A., Zimmer, C. and Wohlschläger, A. M. (2007), 'Selective changes of resting-state networks in individuals at risk for Alzheimer's disease.', *Proceedings of the National Academy of Sciences of the United States of America* **104**(47), 18760–5.

Sporns, O. (2011), 'The human connectome: a complex network.', *Annals of the New York Academy of Sciences* **1224**(1), 109–25.

Stevens, A. A., Tappon, S. C., Garg, A. and Fair, D. A. (2012), 'Functional brain network modularity captures inter- and intra-individual variation in working memory capacity.', *PloS one* **7**(1), e30468.

Strother, S. C., Anderson, J., Hansen, L. K., Kjems, U., Kustra, R., Sidtis, J., Frutiger, S., Muley, S., LaConte, S. and Rottenberg, D. (2002), 'The quantitative evaluation of functional neuroimaging experiments: the NPAIRS data analysis framework.', *NeuroImage* **15**(4), 747–71.

Tagliazucchi, E., von Wegner, F., Morzelewski, A., Brodbeck, V., Borisov, S., Jahnke, K. and Laufs, H. (2013), 'Large-scale brain functional modularity is reflected in slow electroencephalographic rhythms across the human non-rapid eye movement sleep cycle.', *NeuroImage* **70**, 327–39.

Teh, Y. W. (2010), Dirichlet Process, Technical report.
URL: *http://www.gatsby.ucl.ac.uk/∼ywteh/research/npbayes/dp.pdf*

Teipel, S. J., Bokde, A. L. W., Meindl, T., Amaro, E., Soldner, J., Reiser, M. F., Herpertz, S. C., Möller, H.-J. and Hampel, H. (2010), 'White matter microstructure underlying default mode network connectivity in the human brain.', *NeuroImage* **49**(3), 2021–32.

Thirion, B. and Faugeras, O. (2003), 'Dynamical components analysis of fMRI data through kernel PCA', *NeuroImage* **20**(1), 34–49.

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), 273–89.

Varon, C., Alzate, C. and Suykens, J. A. K. (2013), Noise Level Estimation for Model Selection in Kernel PCA Denoising, Technical report.
URL: *ftp://ftp.esat.kuleuven.be/pub/SISTA/cvaron/13-172.pdf*

Waxman, S. G. (2006), 'Axonal conduction and injury in multiple sclerosis: the role of sodium channels.', *Nat Rev Neurosci* **7**(12), 932–41.

Wu, T., Long, X., Zang, Y., Wang, L., Hallett, M., Li, K. and Chan, P. (2009), 'Regional homogeneity changes in patients with Parkinson's disease.', *Human brain mapping* **30**(5), 1502–10.

Wu, T., Wang, L., Chen, Y., Zhao, C., Li, K. and Chan, P. (2009), 'Changes of functional connectivity of the motor network in the resting state in Parkinson's disease.', *Neuroscience letters* **460**(1), 6–10.

Xu, Z., Tresp, V., Yu, K. and Kriegel, H. (2006), Infinite hidden relational models, *in* 'In Proceedings of the 22nd International Conference on Uncertainity in Artificial Intelligence (UAI)', Citeseer.

Zalesky, A., Fornito, A. and Bullmore, E. (2012), 'On the use of correlation as a measure of network connectivity', *Neuroimage* **60**, 2096–2106.

Zang, Y., Jiang, T., Lu, Y., He, Y. and Tian, L. (2004), 'Regional homogeneity approach to fMRI data analysis.', *Neuroimage* **22**, 394–400.

Zuo, X.-N., Di Martino, A., Kelly, C., Shehzad, Z. E., Gee, D. G., Klein, D. F., Castellanos, F. X., Biswal, B. B. and Milham, M. P. (2010), 'The oscillating brain: complex and reliable.', *NeuroImage* **49**(2), 1432–45.