

Hlinka, Jaroslav (2010) Studying spontaneous brain activity with neuroimaging methods and mathematical modelling. PhD thesis, University of Nottingham.

Access from the University of Nottingham repository:

<http://eprints.nottingham.ac.uk/11606/2/HlinkaPhDThesis.pdf>

Copyright and reuse:

The Nottingham ePrints service makes this work by researchers of the University of Nottingham available open access under the following conditions.

This article is made available under the University of Nottingham End User licence and may be reused according to the conditions of the licence. For more details see:
http://eprints.nottingham.ac.uk/end_user_agreement.pdf

For more information, please contact eprints@nottingham.ac.uk

**STUDYING SPONTANEOUS BRAIN ACTIVITY
WITH NEUROIMAGING METHODS AND
MATHEMATICAL MODELLING**

Ing. Mgr. JAROSLAV HLINKA

Thesis submitted to the University of Nottingham
for the degree of Doctor of Philosophy

DECEMBER 2010

Abstract

The study of spontaneous brain activity using functional Magnetic Resonance Imaging (fMRI) is a relatively young and rapidly developing field born in the mid-nineties. So far, sufficiently solid foundations have been established, mainly in validating the neuronal origin of a significant component of observed low-frequency fluctuations in the ‘resting state’ fMRI signal. Nevertheless, the field is still facing several major challenges. This thesis first reviews the current state of knowledge and subsequently proceeds to present original research results that are directed towards overcoming these challenges.

The first challenge stems from the indirect nature of the fMRI recordings, obscuring the interpretation in terms of the underlying neuronal activity. Two investigations related to this are presented. First, I show that increased head-movement, epiphenomenal to altered states of consciousness, can lead to spurious increases in low-frequency fluctuations in fMRI signal. This may adversely affect inferences on the underlying neurophysiological processes. Second, I demonstrate a direct electrophysiological correlate of increased synchronisation of fMRI activity in areas of the much studied default-mode network. By directly studying electrophysiological correlates of fMRI-based functional connectivity, this study took a pioneering approach to confirming the biological validity of the fMRI functional connectivity concept.

Another widely debated question within the field is the optimal method for extracting relevant information from the extreme volumes of neuroimaging data. I present an investigation providing insights and practical recommendations for this question, based on assessing the interdependence information neglected by the commonly used linear correlation for fMRI functional connectivity studies. The results suggest that in typical resting state data, the nonlinear contributions to instantaneous connectivity are negligible.

The third major challenge of the field is the integration of the experimental

evidence into theoretical models of spontaneous brain activity. In the last part of this thesis, such models are discussed in detail, focusing on the two crucial features of observed spontaneous brain activity: functional connectivity and low-frequency fluctuations. Two specific mechanisms for emergence of the latter are proposed, depending either on the local synchronisation dynamics or the regulatory action of particular neuromodulators.

The thesis concludes with discussion of the questions arising from the presented results in the context of the most recent development in the wider field.

Acknowledgements

There are many to whom I owe my thanks for support in the work that led to completion of this thesis. At the first place, I would like to thank my supervisors for their patient guidance throughout my studies. To my principal supervisor Dorothee Auer I am most grateful for the introduction to the world of clinical neuroimaging research I received through being a member of her team at the Division of Academic Radiology. I also highly value her open scientific mind and the trust she showed in supporting me in my plans to combine the neuroimaging work with mathematical modelling.

When I dived deeper into the latter field, Steve Coombes became my patient and enthusiastic guide. I am grateful for the many discussions we had both before and after he finally also became my formal supervisor, the warm welcome to the School of Mathematics and all the invaluable support in my work.

Before I move further, I want to acknowledge the generous funding from the European Union via the Marie Curie Fellowship that allowed me to come to do research in Nottingham. And coming to Nottingham would be much harder, were it not for the AR crowd that I met in the ‘cave’ office in the heart of the QMC hospital. Deprived of sunlight, together we managed to keep our spirits high (enough) and hope unbroken – with the help of cave-dark humour, sports and just sticking together both in and out of the AR. My special thanks go to those that I spent the most time with - Josef, Antonio, Nara, Phili, Maryam, Muftah and Becky. But I cannot forget the others either, particularly the previous generation that introduced us to neuroimaging and life in AR ... and then left Nottingham for some more sunny places like Cyprus or Barcelona. Thanks, Harris, Christine, Ana!

I also want to express my gratitude to my ‘adoptive’ family at Maths – Jonathan, Ioannis, Margarita, Carl, Helmut, Ana, Nikola, Rüdiger and Steve, of course. I thank you not only for scientific discussions and occasional help, but foremost

for the atmosphere and enthusiasm making the research real fun. I wish I could have attended more of the conferences, Friday meetings and pub socials with you guys.

My big thanks go to my new boss Milan for his understanding approach to my thesis-writing obligations especially during the final stage. Also I am grateful to him as well as my colleagues Martin and David for both moral support and discussions in which I shaped my understanding of nongaussianity/nonlinearity in fMRI.

I am immensely grateful that I could keep in touch with my friends back in Prague, helping me to keep continuity with my past and generally the life outside the PhD – to Honza, Filip, Tomáš, Aleš, Petra and many others whom I cannot name all here. During my blitz visits back home, they all found time for sharing something I really like – a long chat ideally with a pot of good tea.

Last but not least, I would like to express my greatest gratitude to my family. They supported me in my studies both by giving hope and advice at the harder times and by just being there as a stable point in my life at all times. Without them, I would not be who and where I am, including the love of puzzles, math and scientific discovery.

Contents

List of Abbreviations	ix
1 Preface	1
1.1 Motivation	1
1.2 Thesis outline	3
I Background	7
2 Investigating the brain	8
2.1 Brain architecture	8
2.2 Neuroimaging methods	11
3 Spontaneous brain activity	20
3.1 Introduction to resting state research	20
3.2 Default mode of brain function and spontaneous brain activity .	21
3.3 Methods for spontaneous brain activity analysis	26
3.4 Alterations in spontaneous brain activity	35
3.5 Neurophysiological underpinnings	38
3.6 What might spontaneous brain activity represent?	39
II Analysing Spontaneous Brain Activity	41
4 Motion as a confound in resting state studies	42
4.1 Nonneural sources of LFF	42

4.2	Motion mediated LFF increase	45
4.3	Summary	58
5	Is linear correlation suitable for fMRI FC?	59
5.1	Introduction	59
5.2	Material and Methods	62
5.3	Results	66
5.4	Discussion	68
5.5	Summary	71
6	EEG underpinnings of fMRI functional connectivity	73
6.1	Material and methods	73
6.2	Results	78
6.3	Discussion	82
6.4	Summary	86
7	Further data-driven investigations	88
7.1	Does motion mediate the link between EEG and DMN FC? . . .	88
7.2	Could motion mediate sedation-induced changes in functional connectivity?	93
7.3	Summary	96
III	Modelling Spontaneous Brain Activity	99
8	Models in computational neuroscience	100
8.1	Hodgkin-Huxley model	101
8.2	Morris-Lecar model	103
8.3	Wilson-Cowan model	106
8.4	Networks of weakly coupled oscillators	112
8.5	Models of spontaneous brain activity	117
9	Exploration of a large-scale interaction model	123

CONTENTS

9.1	Introduction	123
9.2	Model description	124
9.3	Relation between structural and functional connectivity	126
9.4	Modelling LFF	133
9.5	Summary	136
10	Local model of LFF	139
10.1	Introduction	139
10.2	Model description	140
10.3	Results	143
10.4	Discussion	152
10.5	Summary	153
11	Discussion	154
A	Appendix - Selected source code	160
	Bibliography	168

List of Abbreviations

BOLD	Blood Oxygen Level-Dependent
CBF	Cerebral Blood Flow
CSF	Cerebrospinal Fluid
DAN	Dorsal Attention Network
DCM	Dynamic Causal Modelling
DISE	Depolarisation-Induced Suppression of Excitation
DISI	Depolarisation-Induced Suppression of Inhibition
DMN	Default Mode Network
DSI	Diffusion Spectrum Imaging
DV	Dependent Variable
DWI	Diffusion Weighted Imaging
EEG	Electroencephalography
EPI	Echo Planar Imaging
FC	Functional Connectivity
fMRI	functional Magnetic Resonance Imaging
FSL	FMRIB Software Library
FT	Fourier Transform
HB	Hopf Bifurcation
ICA	Independent Component Analysis

LIST OF ABBREVIATIONS

ISI	Inter-Spike Interval
IV	Independent Variable
LFF	Low-Frequency Fluctuations
LFP	Local Field Potential
MEG	Magnetoencephalography
MI	Mutual Information
ML	Morris-Lecar (neuron model)
MNI	Montreal Neurological Institute (anatomical brain template)
mPFC	medial Prefrontal Cortex
MR	Magnetic Resonance
MRI	Magnetic Resonance Imaging
PBVP	Periodic Boundary Value Problem
PCC	Posterior Cingulate Cortex
PET	Positron Emission Tomography
PRC	Phase Response Curve
ROI	Region Of Interest
RSN	Resting State Network
SBA	Spontaneous Brain Activity
SC	Structural Connectivity
SCA	Seed-based Correlation Analysis
SEF 95	95% Spectral Edge Frequency
SIT	Stimulus-Independent Thought
SNIC	Saddle Node on Invariant Circle (bifurcation)
SPM	Statistical Parametric Mapping
TE	Transfer Entropy

Preface

1.1 Motivation

The workings of the human mind have captivated the imagination of philosophers, scientists and probably all curious individuals since the prehistoric times. From this perspective just relatively recent technological advances have facilitated discoveries about its relation to the workings of the brain, which led to the marriage of the Philosophy's daughter Psychology with Neurobiology and natural sciences in general. While this marriage has not always been without disputes, the abundance of interesting results in the field confirms its fertility. Since then the question of mind has become increasingly linked to or even substituted by the question of brain.

In its beginnings, the mapping of human mind-brain relation was dependent on relatively scarce and anecdotal evidence from patients with localised brain damage, invasive experiments and interesting but technically limited EEG findings. The entry of the brain imaging methods and subsequently the functional brain imaging about twenty years ago has had revolutionary consequences. Noninvasive mapping of brain function is now relatively easily available for numerous research groups around the world, leading to an expansion of specialised neurocognitive studies.

From the onset, the field of functional neuroimaging was dominated by a research strategy that views the brain as primarily reflexive, driven by momentarily demands of the environment. Nevertheless, during the last fifteen years there has been a sustained trend of increasing interest in the neuroimaging community that corresponds to another view of the brain workings. This pictures the brain operation as predominantly intrinsic, with adaptive mechanisms of information

collection, maintenance and processing driven mainly internally. Corresponding to the increased support of this perspective, the analysis of spatiotemporal patterns of brain activity in the so-called resting state condition has developed from a controversial niche into one of the main streams of neuroimaging.

Apart from the contribution of maybe less deeply grounded motivations such as availability of a simple study design and the outward attractiveness of the topic, there are principled incentives for paying attention to the spontaneous brain activity research. When one sets aside the evident value these investigations can have directly for the clinical practice, there is also an important promise for the general study of brain–mind relation and function.

Most importantly, taking into account the spontaneous brain activity finally allows the formation of a holistic picture of brain function. Without a significant and detrimental reduction, the brain can be hardly considered as a purely reactive system. Rather, reactions to external stimulation are embedded in and catalysed by its rich intrinsic dynamics. These internal dynamics have been shown to be flexible and subject-dependent, but also relatively robust with respect to brain state. Indeed, it has been shown that taking into account the spontaneous brain activity is relevant for interpretation of responses to external stimulation. Moreover, spontaneous brain activity is becoming increasingly considered as interesting on its own, representing adaptive functional processes of a crucial role for the healthy function and survival of the organism.

While holding the above promises, spontaneous brain activity research comes across with specific challenges: The first of these is shared with functional magnetic resonance imaging in general and relates to the indirect nature of the brain activity measurement method. This calls for improved treatment of systematic noise also known as imaging artifacts. Another method of tackling this challenge is to investigate the link between the functional magnetic resonance imaging results and complementary brain activity measurements such as electroencephalography.

Second, neuroimaging presents us with data of extreme volume and dimensionality. While computing capacity is becoming less of a problem with the improvements in technology, the selection and development of appropriate statistical tools and interpretation frameworks for the optimal utilization of the information in the sheer abundance of data is another major challenge of the field.

Finally, systematic research and integration of the findings of the many studies into a coherent theory of spontaneous brain activity or brain activity in general

is currently limited. To a large extent, this is due to the lack of systematic development of brain system function theory, in particular represented by theoretical and computational models of large scale spontaneous brain activity.

In this thesis, the above presented results, promises and challenges in the field of spontaneous brain dynamics are discussed in detail. In particular, apart from reviewing the work of other groups, several original results of work that was directed towards the main challenges of the field as outlined above are presented.

1.2 Thesis outline

The thesis is organised in the following chapters:

Chapter 2

This chapter sets the stage for the presentation of the field of spontaneous brain activity, as studied using resting state condition neuroimaging data. The basics about brain structure and function and the neuroimaging methods are reviewed. More attention is paid to structural features and specific methods that are especially relevant for the research in the field, or for the original research work presented in this thesis.

Chapter 3

Spontaneous brain activity is introduced in this chapter. After a review of the development of the field in the last fifteen years the quantitative methods for the extraction of important features of spontaneous brain activity are discussed. Further, the key findings in the area of spontaneous brain activity alterations in disease and changed brain state are summarized. Finally, I overview the state-of-the-art knowledge of the neurophysiological underpinnings of the large-scale fMRI resting state features and conclude with some suggestions on the role of the spontaneous brain activity.

Chapter 4

In this chapter, I turn towards the challenges of analysis of spontaneous brain activity data, in particular those acquired by functional Magnetic Resonance Imaging (fMRI). Previously, the power of low-frequency fluctuations in the rest-

ing state networks has been reported to change in altered brain state conditions such as sedation or sleep. Crucially, the neuronal activity patterns are not the only potential source of the fMRI signal fluctuation that might be sensitive to such experimental brain state manipulation. Here, I set out to test whether an increase in one of the potential fMRI measurement confounds, namely the amount of involuntary head-motion, might account for the signal fluctuation increases reported in such experiments.

Chapter 5

High dimensionality and volumes of the neuroimaging data poses a major challenge to interpretation of the spatiotemporal patterns of spontaneous brain activity. Functional connectivity analysis is a way to summarise the complex pattern in simple indices. Nevertheless, the use of linear correlation between time series as a functional connectivity measure has been recently challenged and the need to use nonlinear measures proposed. In this chapter, I both quantitatively and qualitatively probe the suitability of the linear correlation as a tool for functional connectivity estimation. In particular, I use mutual information, a general measure of bivariate dependence, to assess the deviation of standard fMRI data from bivariate Gaussianity, which is a crucial condition for unbiased interpretation of linear correlation.

Chapter 6

Functional connectivity, particularly within the main resting state networks, is a key concept in the current discourse of spontaneous brain activity. Recently, strength of the functional connectivity within the default mode network has been suggested to relate to the level of subject's consciousness. The biological significance of functional connectivity, however, is still not well understood. In this chapter I present an inter-subject comparison of the electrophysiological correlates of the default mode network functional connectivity, uncovering a network-specific and functionally meaningful electrophysiological signature of decreased default mode network functional connectivity.

Chapter 7

This chapter further extends the investigation of the previous chapter via a critical investigation of the potential role of head-movement in the observed elec-

trophysiological signature of the default mode network functional connectivity. Further, I follow the recent studies reporting changes of fMRI FC and ask, similarly as in chapter 4, whether these results could be attributed to changes in the subject's propensity to head-movement in the sedation condition.

Chapter 8

The final challenge addressed in this thesis is the lack of sufficiently developed models for spontaneous large-scale brain activity. Within this chapter, I first review some basic neuronal models used in computational neuroscience. I also include some tools used for modelling neural populations and interacting neural oscillator networks. Finally, I give a review of computational models recently applied to explain some of the phenomena observed by the experimentalists in the field of spontaneous brain activity. Above all, I focus on the accounts of functional connectivity patterns and low-frequency activity fluctuation. The model architectures and explanatory potential are critically discussed.

Chapter 9

Following the introduction given in the previous chapter, I use a computational model consisting of 47 neuronal populations connected according to an anatomical pattern collected from macaque studies to document some of the issues with modelling spontaneous brain activity. First, using our simulations, I discuss the difficulty of predicting functional connectivity pattern from structural connectivity matrix – in particular I focus on the dependence of the structure/function agreement on the parameter settings. Second, I show how realistic low-frequency fluctuation patterns can be modelled in a simple neural population model. Contrary to some previous accounts, this is feasible without inclusion of noise and delays in the model. Importantly, local synchronisation changes play a role in this mechanism, suggesting potential unsuitability of the modelling approach ‘one oscillator per cortical area’ commonly adopted in the field.

Chapter 10

Another potential mechanism for emergence of low-frequency fluctuation is discussed in this chapter. Here I use a localised neural network model that implements the regulatory action of endocannabinoids in a network of synaptically connected neurons. After an initial analysis of behaviour of the network for

weak coupling, I show that for sufficiently strong coupling I obtain nested fast and slow to ultra-slow oscillations. Again, this behaviour robustly appears even in the absence of cortico-cortical transmission delays and noise. I hypothesise that this or a similar regulatory mechanism might play a role in the emergence of the low-frequency oscillations observed in brain imaging experiments.

Chapter 11

The final chapter discusses the questions open by the presented results in the context of the recent developments in the field. The potential future extensions of the research are also outlined.

Following the stylistic tradition, I generally use the plural form of the first person “we” in the body of the thesis. To clarify, the work described in this thesis represents the independent research work I have carried out under the supervision of Prof. Dorothee Auer and Prof. Steve Coombes during my postgraduate study registration with the University of Nottingham. However, at this point I must acknowledge the specific contribution of two of my former colleagues from the Division of Academic Radiology, University of Nottingham. In particular, for chapters 4,5,6 and 7 I have used a resting state fMRI and EEG dataset that was originally acquired by Dr. Charilaos Alexakis and Dr. Ana Diukova. The latter colleague also carried out the standard part of the EEG data preprocessing.

The original research contributions of this thesis have been for the most part already published in peer-reviewed journals. In particular, the work of chapter 4 was published in *Magnetic Resonance Materials in Physics, Biology and Medicine (MAGMA)* [79], the work presented in chapter 6 was published in *NeuroImage* [78] and the work presented in chapter 10 was published in *Physical Review Letters* [80]. Finally, the framework presented in chapter 5, applied to another dataset, has also been recently accepted for publication in *NeuroImage* [81].

Part I

Background

Investigating the brain

Before we give an introduction to the specific subject of this thesis, that is the large-scale spatiotemporal patterns of spontaneous brain activity, we gather in this chapter some important basic facts about the brain and some of the methods used to study it. This is in no way meant to be a comprehensive review. Rather, we set the general stage for presentation of the state of the art in the study of spontaneous brain activity and highlight the points that are most relevant for the original research contributions presented in the later chapters of the thesis.

2.1 Brain architecture

Although most of the basic information on brain anatomy and function presented further can be considered common knowledge, we generally account to Kingsley's neuroscience textbook [95]. We refer the reader to this or other comprehensive textbooks such as [94] for more detail.

2.1.1 Brain cells and signalling

The central nervous system is composed of two types of cells (*neurons* and *glia*), enclosed by three layers of connective tissue that form the *meninges*. Neurons are composed of a *soma*, an *axon* and *dendrites*, where the axon and the dendrites are processes that extend from the soma (cell body). See a schematic of a neuron at Figure 2.1. In the central nervous system, collection of somas is commonly called a *nucleus* and collection of collated axons a *tract*. The central nervous system consists of the *brain* and the *spinal cord*. The brain is further divided into the *forebrain* (including in particular the *cerebral hemispheres*, *basal ganglia* and the *thalamus*), the *brainstem* (consisting of the *midbrain*, the *pons* and the

medulla) and the *cerebellum*.

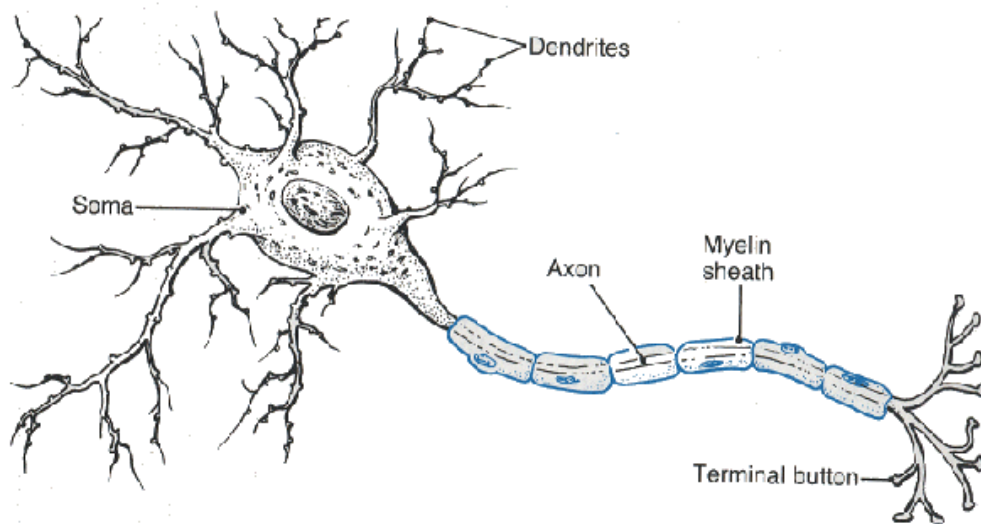


Figure 2.1: A schematic of the neuron structure. Adopted from [26].

There are two important means of signalling in the nervous system: electrical and chemical. Generally, electrical signalling is used for intracellular signalling, while chemical signalling is used for intercellular transmission. Ion imbalances across the neuronal cell membrane lead to a negative electrical potential inside the cell and enable electrical signalling.

To understand the signalling process, it is useful to consider the microanatomy of the axon. The following three regions can be defined within an axon: the *initial segment*, the *axon proper* and the *terminal bouton*. Generally, the initial segment is specialized for initiation of the action potential upon suprathreshold stimulation, the axon proper for its effective conduction due to myelination and the terminal bouton contains specialised structures for chemical communication with other cells.

Upon initiation, action potentials are spread along the axon due to combination of passive conduction between *nodes of Ranvier* (regularly spaced gaps in the myelin sheath around an axon or nerve fibre) and successive activation of voltage-gated ion channels in the nodes of Ranvier, which leads to rapid hyperpolarisation. At the terminal bouton, *synaptic vesicles* release neurotransmitter chemicals into the *synaptic cleft*. The neurotransmitter binds to specific locations of the postsynaptic membrane and activates ion-channel (or second-messenger cascades). The *depolarisation* (in the case of an excitatory synapse) or *hyper-*

polarisation (in the case of an inhibitory synapse) spreads passively along the dendrite or soma as a local current. The resulting signal at the initial segment of an axon of the receiving neuron is the sum of all the reaching local currents.

2.1.2 Cortical structure and connectivity

The structure of brain tissue and connections is quite stable across individuals of one species. In the following, we shall focus only on describing some details of the structure of the human cerebral cortex and thalamus and their connections.

To a first approximation the thalamus serves as a relay station for the sensory information going towards the cortex and motor information coming out of the cortex, having reciprocal connections to the cerebral cortex. It can be divided into several groups of nuclei. Some of the nuclei are connected to very small parts of the cerebral cortex; these are so-called specific thalamic nuclei, while other nuclei are non-specific. An important group is the reticular thalamic group, sending inhibitory connections to other thalamic nuclei and playing a role in the wake-sleep cycle [170]. The sensory and motor representation in both thalamus and cortex is *somatotopic* (maintaining spatial organisation of the body).

The cerebral cortex is composed of 2-4 millimetres thick layer of *grey matter* and an underlying mass of *white matter*. The grey matter is organised into 6 *cortical layers*, defined by histological criteria such as size and number of various cells. The principal cell type in the cerebral cortex is the *pyramidal cell*. Pyramidal cells are found in all but the first layer and have an apical dendrite extending from the peak of the triangular shape of the soma and reaching the surface layer, with basal dendrites spreading horizontally and a single axon projecting into the white matter. The pyramidal cells are the only cortical neurons projecting outside the grey matter, all others are *interneurons*. The cortex input is as follows:

- association fibres from other parts of the cortex synapse in all layers within one cortical column
- fibres from specific thalamic nuclei synapse primarily in layer IV within one cortical column
- fibres from the non-specific thalamic nuclei and from the specialised sub-cortical nuclei synapse in all layers, terminating commonly in large areas after branching.

2.2 Neuroimaging methods

The aim of this section is to give an overview of neuroimaging methods which are available for studying spontaneous brain activity, with a specific focus on those which are used or referred to in this thesis.

Neuroimaging is a rapidly developing field, offering a range of methods. These can be divided into structural (anatomical) and functional neuroimaging methods. Major structural methods are Computer Tomography (CT) and Magnetic Resonance Imaging (MRI), introduced in the late seventies and offering 2D or 3D information about anatomical structure of brain or other tissues. These methods, particularly MRI, have undergone profound development since being invented. The functional imaging methods include Positron Emission Tomography (PET) and functional Magnetic Resonance Imaging (fMRI) [136].

There is a slightly separate family of methods commonly also included in the neuroimaging methods, which are based on direct measurement of neuronal electromagnetic activity. The noninvasive representatives include particularly Electroencephalography (EEG) and Magnetoencephalography (MEG). On the other hand, invasive recordings from subdural electrodes are utilised, mainly for clinical purposes, in Electrocorticography (ECoG). The main advantages of this family of methods are the high temporal resolution (in order of milliseconds, see Figure 2.2 adapted from [32] for comparison) and direct sampling of electrical neuronal activity rather than of the metabolic correlates as in the case of fMRI (to some extent, this may be considered a disadvantage as well, if we are more interested in the metabolic demands). The main disadvantage of noninvasive electrophysiological methods lies in poor spatial resolution and poor localization of sources and preferential sensitivity of these methods to specifically located and oriented neurons.

In the following, we briefly describe the principles, advantages and drawbacks of the methods most relevant for this thesis.

2.2.1 MRI

Introduction to MRI

The use of the principles of nuclear magnetic resonance for medical scans of the brain and other organs was introduced in the during the 1970s and 1980s. Although the methods have developed in great detail, the main principles are

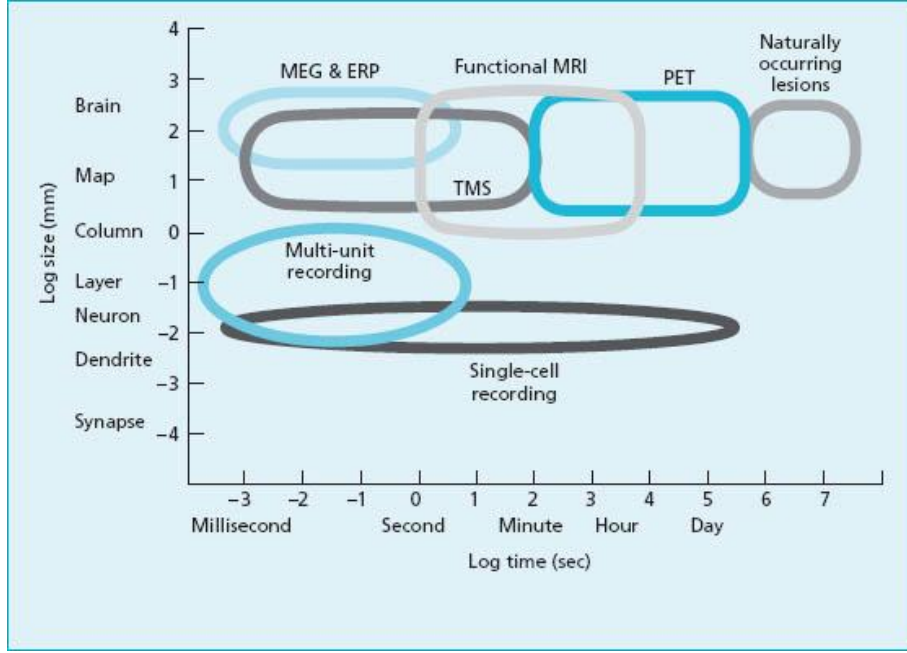


Figure 2.2: The methods of cognitive neuroscience can be categorized according to their spatial and temporal resolution. Adopted from [32].

reasonably straightforward to be covered in brief in this section.

The basic fact the method relies on is that the atoms in the tissue behave like tiny magnets, due to the spins of the nuclei (it suffices for our purposes to only consider hydrogen atoms containing a single proton in the nucleus). When a strong external magnetic field is applied (conventional medical scanners have fields of 1.5 or 3 Tesla, which is about 100000 times stronger than the Earth's magnetic field), the spins align in parallel or antiparallel with the applied field, with a slight prevalence of the parallel orientation (to the order of 0.0001%). This leads to a net steady state magnetization B_1 . In more detail, each spin precesses around the B axes with a so-called Larmor frequency

$$\omega = \lambda B,$$

where B is the applied external magnetic field and λ is the gyromagnetic ratio, with a value of 42.576 MHz/Tesla for 1H nucleus. When an electromagnetic pulse with Larmor frequency (called *RF pulse* for radiofrequency) is applied to the sample, the spins excite, the axis of precession changes and the phases of the spins synchronise. After switching the RF pulse off, the spins re-emit energy in the form of an RF signal of the Larmor frequency whilst returning to the baseline precession state. During the read-out phase, the scanner receives the

re-emitted signal and further processing is applied to this signal to recover the spatial and structural information it conveys about the excited sample.

Note that all the spins precess and also emit RF radiation on the same frequency (and in phase, since initially brought in coherence by the excitation pulse), unless spatial inhomogeneity of the magnetic field occurs or is imposed. In practice, small magnetic fields with a linear intensity gradient along a chosen spatial direction (commonly called magnetic gradients and gradient direction) are used to introduce heterogeneity of Larmor frequencies across the sample, perpendicular to the gradient direction so as to distinguish signals re-emitted from different locations.

The return to the baseline state consists of two main processes. The first is the dephasing of the spins with respect to each other. This is an exponential process governed by the so-called $T2$ constant, which is further pronounced in case of inhomogeneity of magnetic field due to fine tissue structure – when these effects are included, the temporal constant is then called $T2^*$. The second process taking place is the longitudinal recovery of the net magnetization, an exponential process governed by a $T1$ constant. Both these processes contribute to the rapid decay of the re-emitted signal after switching of the RF pulse. Since different tissues have different $T1$ and $T2$ values, manipulation with the scanning sequence (excitation, readout) can make the readout signal more susceptible to $T1$ or $T2$ processes. Thus, we are able to recover images with various contrasts distinguishing various types of tissue (grey matter, white matter, cerebro-spinal fluid, lesions, or more oxygenated tissue as utilised in functional MRI techniques discussed later). The necessary step to do that is to be able to decode spatial information from the signal to recover spatial origin of the signal within the sample. To achieve this a combination of three methods is commonly used, namely:

- slice selection (by applying a small gradient during application of narrow-band RF pulse)
- frequency encoding (applying small gradients during read-out of the signal)
- phase encoding (applying small gradients between excitation and readout to dephase the spins differentially in different locations of the sample across the direction of the gradient in a controlled way)

While these principles generally cover the processes needed to acquire anatomical images sensitive to tissue type, we refer the reader to a comprehensive book [76]

for a detailed account of the underlying physical principles and scanning sequence design.

fMRI

The basic idea of fMRI is to capture the (brain) activity in consecutive temporal windows to be able to study the dynamics of the spontaneous activity or the response to an experimental stimulus. This firstly imposes some demands on the speed of acquisition, and secondly calls for a contrast that is sensitive to changes of brain activity.

The first step of enabling rapid acquisition was solved in the early eighties by Sir Peter Mansfield in proposing an imaging sequence called Echo-Planar Imaging (EPI) [117]. In the recent implementations, it offers whole brain coverage with few millimetre resolution acquired in a couple of seconds. For example, datasets with $64 \times 64 \times 35$ cubic voxels of $3.25 \times 3.25 \times 3.0$ mm acquired every 2.1 seconds were used in this thesis, see Chapter 4 for details of the methods).

Second, a specific contrast was observed in the early nineties to be sensitive to haemodynamic response to neural activity and gave rise to the so-called Blood Oxygenation Level Dependent (BOLD) fMRI contrast [136]. We shall cite one of the biggest authorities in the field, Prof. Nikos Logothetis: *“As its name suggests, the BOLD contrast mechanism alters the $T2^*$ parameter mainly through neural activity-dependent changes in the relative concentration of oxygenated and deoxygenated blood. Deoxyhemoglobin (dHb) is paramagnetic and influences the MR signal unlike oxygenated Hb. In the presence of dHb, the $T2$ value decreases quadratically with field strength, as expected from the dynamic averaging owing to diffusion in the presence of field gradients. The effects of dHb on $T2^*$ are even stronger, as first noticed by Ogawa et al. Specifically, Ogawa & Lee observed that blood vessel contrast varied with changes in blood oxygen demand or flow”* [113]. While the exact mechanism of BOLD signal is still not fully understood, since it is dependent on the interplay of many biological factors, the basic principle can be summarized as follows: The increased neuronal activity triggers a pronounced increase in Cerebral Blood Flow (CBF) accompanied by a less pronounced increase in Oxygen Extraction Fraction (OEF). This leads to a decrease in tissue dHb content and conversely to a higher signal intensity.

Regarding the neuronal correlates of the BOLD signal, the cited paper shows results of simultaneous BOLD and electrophysiological recordings, suggesting

that the BOLD response primarily reflects the input and local processing of neuronal information rather than the output signals, which are transmitted to other regions of the brain by the principal neurons [113].

Another important finding further complicating a straightforward interpretation of the fMRI data is that the haemodynamic response function linking the neuronal activity and BOLD signal increase is heterogeneous across the cortex of an individual subject, between subjects and also presumably dependent on other conditions, which makes the precise quantitative interpretation of BOLD signal in terms of neural activity a difficult task.

In this context it is also appropriate to mention briefly another complication which usually comes with fMRI BOLD measurements. This is the problem of structured noise in the MRI data, also known as *imaging artifacts*. Probably the most common and cumbersome are:

- artifacts connected to physical motion of the subject during acquisition
- cardiac and respiratory activity related artifacts
- scanner related artifacts (i.e. scanner drifts, inhomogeneity of signal, susceptibility artifacts)

When mentioning these artifacts, we also outline the basic preprocessing steps which are usually carried out before feeding the data into a statistical analysis:

Reslicing the data Since the acquisition is done slice by slice, not all the data in one 3D volume actually reflect the same moment. The simplest way to correct for this is to use linear interpolation based on information about the slice acquisition order

Motion correction Physical and apparent (scanner-drift related) movement of the head as well as pulsation effects usually leads to spatial misalignment of the 3D volumes. To correct for this effect typically a simple linear transformations consisting of translations and rotations is used, with parameters optimized to minimize some goodness-of-fit measure such as mutual information between the two images

Spatial smoothing is used to raise the signal to noise ratio by averaging out the noise (based on the assumption of spatial independence/non-smoothness of noise and smoothness of signal)

Frequency filtering is often applied to get rid of contributions which are most likely to be physiological artifacts ($> 0.1\text{Hz}$) or a scanner drift or trend ($< 0.01\text{Hz}$). Nevertheless, the use of spectral filtering methods for physiological artifact removal has serious drawbacks. In particular, as commonly above the Nyquist frequency, the physiological artifacts often alias into the pass-band of the frequency filter.

Filtering out physiological signals is an advantageous alternative or complement to the frequency filtering when the simultaneously acquired physiological signals are available. Example of such a method is the RETROICOR algorithm [67]. Here, the physiological signal is expressed by a low-order Fourier series (typically of order two) in terms of the phases of the cardiac and respiratory cycles. Subsequently, the contribution of the estimated physiological signal is filtered out of the image data (for each voxel separately).

MRI diffusion

Another important specialised MRI technique is the diffusion-weighted imaging (DWI). The principle is that additional small gradients along multiple directions are applied in a specific manner, to create a contrast which is sensitive to (water molecule) diffusion along a given direction. Combining multiple diffusion weighted images enables one to recover information about the overall strength of diffusion and its directionality. This in turn enables one to distinguish white and grey matter by means of the overall strength of diffusion (captured by e.g. average diffusivity index, AD) as well as direction and overall anisotropy of diffusion (captured by e.g. Fractional Anisotropy (FA)). Changes in these measures can be sensitive to neurodegenerative disease processes, e.g. in multiple sclerosis [45].

Pushing the limits of methods of DWI, new approaches to high angular resolution diffusion weighted imaging such as Diffusion Spectrum Imaging (DSI) allow good resolution of directionality profiles of white matter tracts in each voxel. This is a promising avenue for solving the problem of interpretation directionality in voxels containing *crossing fibres*. This in turn brings important improvement in diffusion-based *tractography*, a method of extracting the trajectories of brain white matter tracts in vivo using noninvasive methods of MRI. Thus, noninvasive investigation of structural brain connectivity of single human subjects becomes available, of potential interest for clinical applications but also

providing priors for functional modelling studies. A first pioneering example of such an application is [84], but the potential of the structural priors in models is yet far from being properly utilised.

2.2.2 EEG

Compared to MRI, the EEG method is conceptually simple and more direct; nevertheless the interpretation of EEG data in terms of neural activity is also not straightforward. Electroencephalography is the neurophysiologic measurement of the electrical activity of the brain by recording from electrodes placed on the scalp (or, in special cases, subdurally or in the cerebral cortex). The resulting traces are known as an electroencephalogram and represent an electrical signal (mainly postsynaptic potentials, not action potentials [95]) from a large number of neurons. While EEG offers the advantage of millisecond temporal resolution, the spatial resolution is poor in comparison to fMRI. That stems from a number of reasons:

- The limited number of channels acquired (typically several tens to a hundred of electrodes)
- The non-unique solution to the inverse problem of detecting electrical sources inside the head from the surface measurement
- The complex conductance pattern of head which further complicates estimating the solution to the inverse problem.

Due to these reasons, in most studies, researchers do not attempt to recover the exact distribution of the sources inside the brain, rather staying at the surface electrode/channel level, assuming that e.g. the occipitally placed electrodes carry information mainly from visual cortex areas.

Typically, EEG recording is divided into frequency bands for ease of description and analysis. The main four are (note that the behavioural characterization is extremely simplified):

Delta-band (0.5-4Hz) – prevalent in deep sleep

Theta-band (4-8Hz) – widespread in the onset of sleep and in deep relaxation, but also linked to some memory-related cognitive processing

Alpha-band (8-13Hz) – prominent rhythm in a relaxed state, pronounced in visual areas in eyes closed condition, diminishes when eyes opened

Beta-band (13-30Hz) – prominent rhythm during cognitive activity

Aside from these frequency bands, fluctuations in higher frequencies than 30Hz called Gamma-oscillations are extensively discussed in cognitive neuroscience. Quite recently, it has been proposed by Steriade that different brain rhythms are grouped within complex wave-sequences [170]. This perspective may help to approach electrophysiological data more effectively, but has not been widely utilised yet. The straightforward categorisation of behavioural correlates of the brain rhythms is further complicated by their spatial heterogeneity, inter- and intra-subject variability and reactivity to brain state or specific stimulation. For a thorough but readable account of rhythmic activity in the brain we refer the reader to [24].

2.2.3 Other neuroimaging methods

We briefly review some other brain imaging methods which, although not used in our experiments, are highly relevant for the research in spontaneous brain activity. For the sake of brevity we have to exclude other methods such as Near Infrared Spectroscopy and optical imaging methods in general [64], although they are also becoming applied in the field.

MEG

Magnetoencephalography is a recent tool for measuring neural activity of the brain. It uses surface sensors (often called channels) in usually higher density than EEG (of several hundreds sensors) to measure subtle changes of magnetic field. The whole measurement is technically very demanding, inquiring special shielding from magnetic fields from the environment and cooling of super-conductive coils in the detector system. The reward lies in obtaining a better starting point for spatial localization of sources due to the following facts:

- Higher number of detectors than in EEG
- The conductivity does not affect the spreading of magnetic field and therefore the estimation of sources does not rely on so many assumptions and prior knowledge of brain and skull anatomy and physical properties.

Still, as with EEG, the solution to the inverse problem of localisation of the sources from the measurements is not unique [177], and therefore the solution

can be uniquely defined only through introduction of a set of constraints. One approach to recovering the spatial information about the localization of the sources is the family of beamformer methods, an example application of which is Dynamic Imaging of Coherent Sources (DISC) [74]. The main principle of beamformer techniques is in obtaining estimates of electric field time-course at a given location in the brain as a weighted sum of surface measurements, where the weightings are derived from the channel covariance matrix under some rather general assumptions [22].

Invasive electrophysiological recordings

One of the major difficulties of EEG and MEG recordings was the inverse problem of localisation of the sources of the measured signals. With invasive electrophysiological recordings, this problem is largely attenuated. Thus, they serve an important role in validating the hypothesis generated using the more generally applicable neuroimaging methods. The invasive nature of the procedure on the other hand limits the potential use in humans and to great extent in animals as well for both ethical and experimental reasons. This is reflected in the limited availability of invasive electrophysiological recordings. In particular, the several published human studies report results from measurements taken as part of preoperative assessment of patients with pharmaco-resistant epilepsy.

The broad-band signal obtained from microelectrodes can be generally divided by filtering into the *local field potential* (LFP) and the *multi-unit activity*. The former is obtained using a low-pass filter at $\sim 300\text{Hz}$ and corresponds generally to the sum of dendritic synaptic activity within a localised volume of tissue. The multi-unit activity is then the fast component of the signal, corresponding to individual action potentials or spikes, which under good circumstances can be ascribed to individual neurons using *spike-sorting* techniques.

Analysis of invasive electrophysiological recordings in the context of large-scale spontaneous brain activity has led to several interesting observations. For example, interhemispheric correlations of slow spontaneous fluctuations in firing rates and LFP power modulations in human sensory cortex have been directly observed [135], suggesting the role of long-range LFP power synchrony in functional connectivity. Second, LFP power in the default mode network areas has been shown to have the same reactive task-related properties as the fMRI signal, confirming the neural origin of the task-related deactivations in the default mode network.

Spontaneous brain activity

In this chapter we introduce the modern research into large-scale spontaneous brain activity (SBA) using neuroimaging methods, particularly fMRI. First, we describe the development of the field in the last fifteen years. This is followed by an overview of the quantitative methods used for extraction of the most important features of spontaneous brain activity. We also discuss the key findings in SBA changes in disease and altered brain state and offer an overview of the state of the art in understanding the neurophysiological underpinnings of the large-scale fMRI resting state features. We conclude with some speculation on the role of the spontaneous brain activity.

3.1 Introduction to resting state research

Functional neuroimaging (utilising PET and later fMRI) focused at its outset on studies of brain responses to carefully controlled sensory, cognitive and motor tasks. These experiments fit well with the view of the brain as driven by the momentary environmental demands [149]. However, during the last fifteen years the interest in the study of intrinsic, un-driven or on-going brain activity has steadily increased. Given the nature of this activity, the terminology in this area is not yet fully established. In the following we shall refer to *resting state* whenever we focus on the experimental condition during which spontaneous brain activity data are typically acquired, while the term *spontaneous brain activity* is left for denoting the brain activity not directly related to external stimuli manipulation. Note that the term *intrinsic brain activity* was proposed with a similar meaning, suggesting a distinction from *evoked activity* [147]. Although there are some methodological and theoretical challenges in the measurement and interpretation of resting state data (for a sceptical view see e.g. [129]), resting state

research is gaining in significance in a way eliciting an impression of a paradigm shift in neuroimaging [146].

The beginnings of growing interest of neuroimaging researchers in the resting state of brain lie in the late nineties. Probably the earliest, nowadays almost classical observation of functionally meaningful patterns in the resting state fMRI is the study by Biswal et al. [15]. This study demonstrated that even during resting condition, temporal signals from bilateral motor cortices mutually correlate, while this correlation is specific for the motor network. Nevertheless, the rise of the interest in spontaneous brain activity patterns in the first stages was most probably driven mostly by the observations of activity in the so-called *default mode network* (DMN). Mainly for this reason, in the following review of recent spontaneous brain activity research we adopt the perspective of evolution of the concept of the DMN.

3.2 Default mode of brain function and spontaneous brain activity

3.2.1 Deactivation during active task

As stated before, early fMRI studies were primarily directed at detecting increases in signal during carefully selected tasks meant to affect activity in specific brain modules. Nevertheless, these studies also tended to show spatially consistent patterns of *deactivations* across a wide range of task conditions. In 1997, Shulman and his colleagues [161] published a meta-analysis of several studies that reported such consistent BOLD deactivations during attention-demanding tasks in specific areas of cortex (precuneus gyrus, cingulate cortex, medial prefrontal cortex). This finding has been replicated by independent researchers in a number of studies [13, 122]. As well, the results have been generalised across modalities by reporting similar findings using PET imaging [148]. In the latter paper, Raichle hypothesised that there exists a default mode of brain function, during which a set of areas including the precuneus gyrus and medial prefrontal cortex is tonically active, continuously gathering and processing information about the external and internal environment.

3.2.2 Spontaneous fluctuations anticorrelated networks

Another step in the exploration of spontaneous brain activity was the observation that during the resting state condition the default network is actually not *tonically* active, but instead fluctuates at typical frequencies of $0.01 - 0.1\text{Hz}$, as was pointed out by Fransson in 2005 [53]. More detailed analysis in his paper, confirmed also by a concurrent report by Fox et al. [51] revealed that while the default network fluctuates coherently (in-phase) as a whole, there is another set of brain areas that fluctuates out-of-phase and consists of the dorsolateral prefrontal cortex, dorsal premotor cortex, including the supplementary motor area and the inferior frontal gyrus/ventral premotor cortex, bilateral supramarginal gyrus, the posterior parietal cortex, insula and the extrastriate cortex. To be more precise, regression analysis of each voxel timecourse against the precuneus area average timecourse has shown positive correlations for default network areas and negative correlations in the other set of areas. While the precise functional interpretation of the two networks is still debated, the original proposal of Fransson was that “... *the brain recurrently toggles between an introspectively oriented mode (default mode) and a state-of-mind that tentatively might be interpreted as an extrospectively oriented mode that involves a readiness and alertness to changes in the external and internal environment.*” Similar findings led Michael Fox and his colleagues from the Raichle’s group to formulate the hypothesis that: “*The human brain is intrinsically organized into dynamic, anticorrelated functional networks*” [51]. See Figure 3.1 for illustration of the two anticorrelated functional networks as detected by the Raichle’s group.

3.2.3 Resting state in behavioural context: processing load & SIT

Once the phenomenon of coherent resting state fluctuations had been reported by independent groups and quite well established, more attention has been focused on its functional and behavioural significance. To mention just two of these attempts, default mode activity has been studied under sustained memory task [54] and in connection to daydreaming [120].

In the former paper, Fransson describes the decrease in fluctuations and connectivity within DMN under sustained working memory task. He presents three main findings:

- The extent of correlating areas is decreased

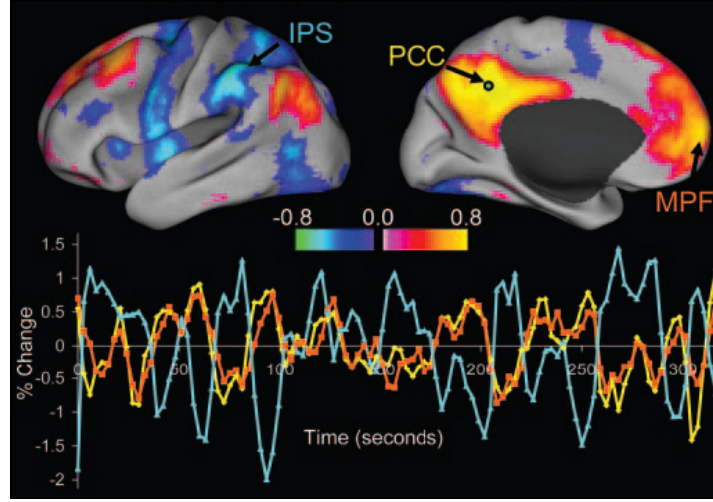


Figure 3.1: Figure shows intrinsic correlations between a seed region in the precuneus cortex and all other voxels in the brain for a single subject during resting fixation. The spatial distribution of correlation coefficients shows both correlations (positive values) and anti-correlations (negative values), thresholded at $R > 0.3$. The time course for a single run is shown for the seed region (precuneus cortex, yellow), a region positively correlated with this seed region in the medial prefrontal cortex (orange), and a region negatively correlated with the seed region in the intraparietal sulcus (blue). Reproduced from [51].

- The magnitude of correlation to precuneus seed is decreased in most DMN areas
- The mean power density of intrinsic low-frequency signal fluctuations decreases under sustained task

Clearly, these three findings can be seen as three manifestations of a general fluctuation decrease. On the other hand, this study also confirmed the observation of a well-defined pattern of spontaneous fluctuations even during a sustained attention-demanding task.

The latter paper by Mason et al. [120] explores the apparent possibility of a link between the observed default network activity and daydreaming activity of the resting subject. During a paradigm validation session, the subjects' were presented practised and novel tasks, and a simple questionnaire was used to assess the number of *stimulus-independent thoughts* (SIT, daydreaming) that occurred during the sessions. The difference was statistically significant. Later, in an epoch paradigm, they were presented both novel and practised tasks, as well as periods of rest. It was found that during the novel tasks the decrease of DMN activity was more pronounced than during the practised tasks (with

respect to the rest) – see Figure 3.2. To further support the hypothesis that the DMN activity is related to daydreaming, the decrease difference was correlated with result of the Imaginal Processes Inventory (IPI) measuring the propensity for daydreaming. The correlation was found to be significantly positive for most of DMN areas. While the possibility of misattribution of DMN activity to SIT instead of *stimulus-oriented thought* (SOT) has been pointed out in an immediately following technical comment published in the same journal [65], the findings seem to be supported by other new studies such as [31].

3.2.4 Alternative analysis methods and the zoo of networks

When viewing the results of the spontaneous brain activity research from a perspective, two interlinked characteristic features of the dynamics stand out. The first is the typical temporal property of *low-frequency fluctuation* of the signal (LFF, 0.01 – 0.1Hz, [53]). The second is the fact, that spatially distant but functionally related regions show coherent activity dynamics, a phenomenon called *functional connectivity* (FC) [56, 59] in the general context of imaging neuroscience. Many of the findings discussed above, particularly the detection of the DMN and its anticorrelated networks by means of regression analysis, were in principle obtained using a simple analysis method called *seed-based correlation analysis* (SCA).

SCA consists of a few elementary steps. First, a suitable seed voxel or area is chosen, which is assumed to lie within a core of a network of interest. Then, the timecourse of the seed is correlated to that of each brain voxel independently, deriving a spatial map of correlations to the seed-voxel. This map then undergoes appropriate thresholding and statistical testing.

One of the main disadvantages of the seed-based correlation analysis is the arbitrariness of the choice of the seed-voxel. Also, due to the principally bivariate analysis, this approach is not able to effectively utilise the multidimensional nature of the data and is very sensitive to the effects of imaging artifacts discussed in section 2.2. Soon after the topic of spontaneous brain activity and resting state data analysis had attracted the attention of the wider neuroscience community, an alternative method called Independent Component Analysis (ICA)[34] has been applied to the resting state fMRI data [97, 10] and become widely used. (Note that the earliest ICA application to fMRI, though to task-related data, dates back probably to 1998 [125].)

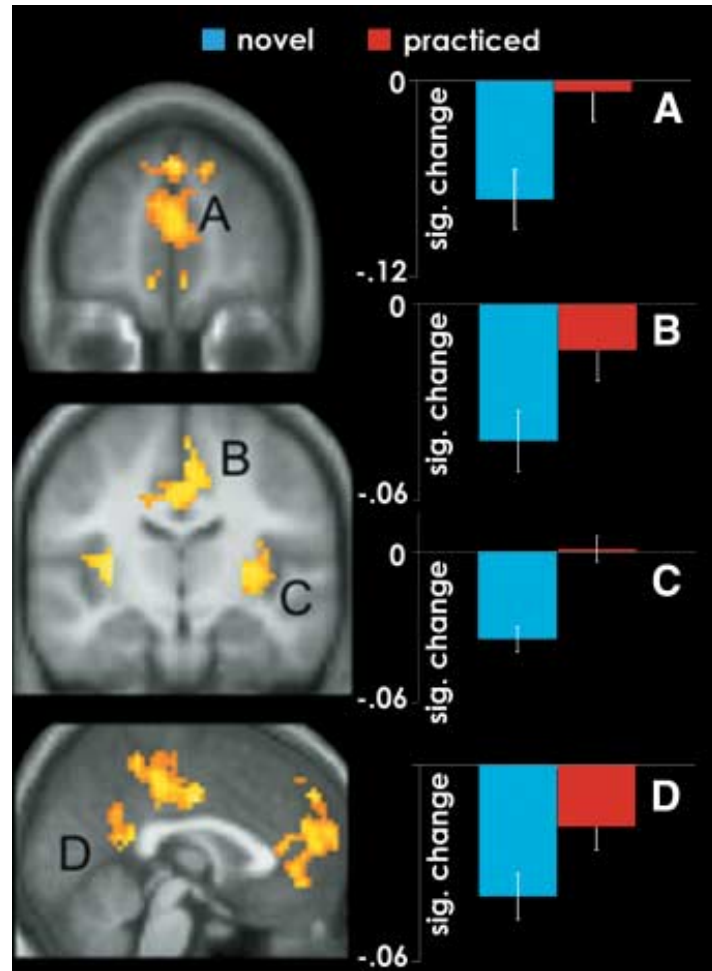


Figure 3.2: Figure shows the BOLD brain activity signal change from resting state to steady state task of two different difficulties: a more difficult novel task and an easier, since already practised, task. Note the prominent effect of task difficulty on the activity change. Note that the tasks in practised and novel task category were matched to avoid intrinsic task difficulty as a confound. Illustration taken from [120].

While the resting state data analysis methods are described in a later section, we note here three critical advantages of ICA over the SCA approach. Firstly, the method is generally model-free – no choice of seed voxel or region is needed. Secondly, the method naturally decomposes the data into a *set* of spatiotemporal components, each representing the location and dynamics of a single network. And thirdly, as the noise is typically classified into separate components, the actual spatial maps of the networks are more robust with respect to noise.

In our view, the use of ICA on resting state data contributed strongly to change of the general view of spontaneous brain activity. While at some stage, the spontaneous brain activity might have been (simplistically) iconified by the DMN, the model-free ICA approach made it much clearer that there are many functional networks, whose activity slowly fluctuates with high level in internal consistency – with all of them fluctuating concurrently. This comes back to the original observations of Biswal [15] of the spontaneous synchronous oscillations in the motor network – the choice of motor network was driven only by convenience, while similar synchronous oscillations happen also in other functional networks such as visual or somatosensory. The name *resting state networks* (RSNs) has been widely accepted for these networks, although they have been shown to correspond to functional networks related to specific cognitive/behavioural activities [165].

To give an example, in 2006, with a particular implementation of ICA, namely Probabilistic Independent Component Analysis (PICA) [10], De Luca et al. [38] described at least 5 RSNs consistent across subjects (see Figure 3.3). Note that while robust identification of only up to ten networks on a single subject level is usually possible, the group studies have been able to detect already several tens of functionally meaningful networks [98], utilising the combined wealth of data.

3.3 Methods for spontaneous brain activity analysis

In this section we introduce some of the statistical methods used for analysis of fMRI spontaneous brain activity recordings.

From mathematical point of view, acquisition of resting state fMRI data leads to a 4D scalar field of intensity values (3 spatial dimensions and 1 temporal). For an example, spatiotemporal resolution of $3.25\text{mm} \times 3.25\text{mm} \times 3.0\text{mm}$ ($\times 2.1$ seconds) used in our data presented in Chapter 4 can be typically reached while keeping a reasonable signal to noise ratio. With this resolution, the images cov-

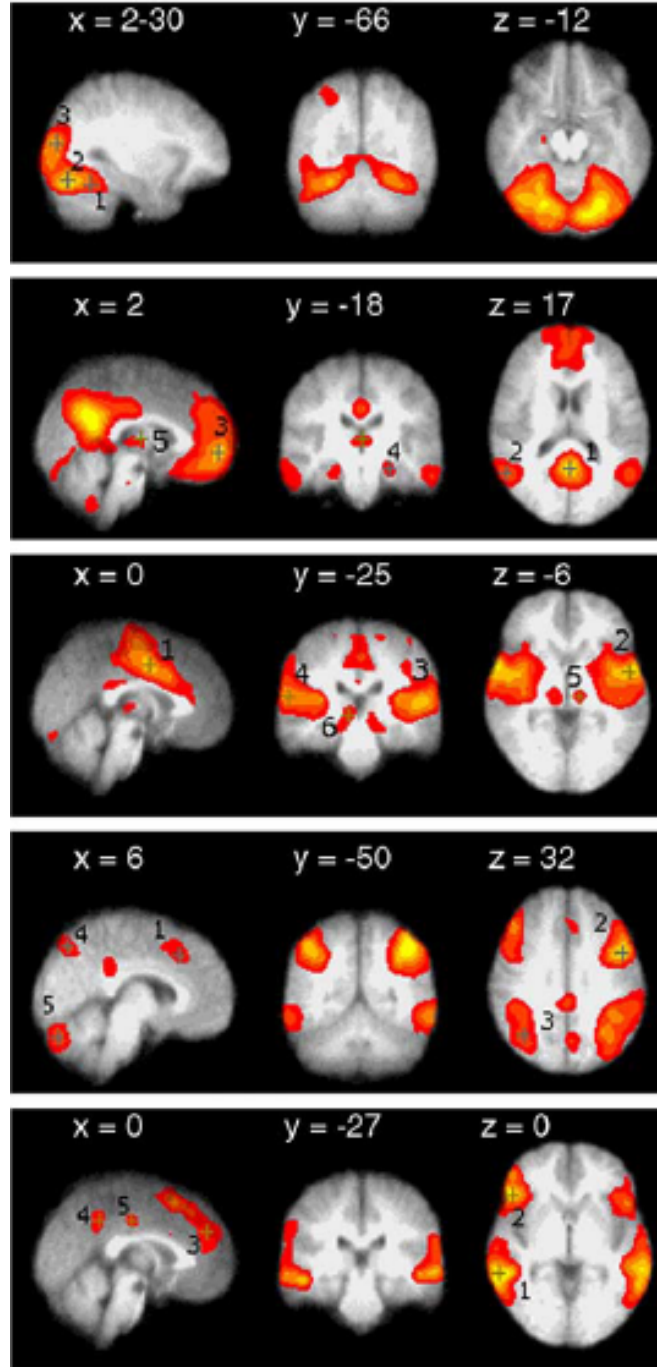


Figure 3.3: Figure shows the location of five resting state networks extracted consistently across subjects using PICA by deLuca et al. [38]. Note the second from top network shown, spatially corresponding to the default mode network described previously. Source: [38].

ering the whole brain have a size of $64 \times 64 \times 35$. Therefore we obtain about 100000 voxels (volume pixels) for each 3D volume. Multiplied by several hundreds of time points in typical resting state scan, we obtain large amounts of data demanding conceptually reasonable as well computationally efficient analysis methods yielding easily visualize-able output. While we have introduced some of the preprocessing steps in the section 2.2, here we discuss the methods used to summarise the statistical relations in the data.

Since the first reports of meaningful spontaneous brain activity patterns, two interlinked features of the data stood out. As mentioned earlier, the first was the low-frequency fluctuation (0.01 – 0.1 Hz) of the signal, while the second is the synchrony of these low-frequency fluctuations among distant areas, called functional connectivity. For quantifying the LFF, simple computation of spectral power in the low-frequency band is commonly used. Alternatives include the use of the variance or coefficient of variation, or specific measures such as fractional Amplitude of Low Frequency Fluctuations (fALFF)[182].

The terms functional connectivity and effective connectivity have been used to describe neural interactions, typically on the level of neural assemblies, even before the application to neuroimaging. In structuring this section, we shall adhere to the distinction made by Friston et al. [56, 59] between the *functional connectivity* as temporal correlations between remote neurophysiologic events and *effective connectivity* as the influence one neural system exerts over another.

3.3.1 Functional connectivity methods

For the purpose of this review, we shall divide the functional connectivity methods of SBA analysis used in the literature to bivariate methods, multivariate methods and ‘complex methods’ which usually include a mix of both bivariate and multivariate methods. The list is far from complete since the range of methods applied is growing quickly as the topic of connectivity attracts more mathematicians and data analysts.

Bivariate methods

Probably the simplest approach to studying functional connectivity between regions of brain is to use linear Pearson’s correlation coefficient of timecourses from a pair of voxels as a measure of strength of connection between the two time series. This simple correlation method can be used to derive a network

of all voxels connected to a given seed voxel (or seed region consisting of several tens to several hundreds voxels and represented by the average timecourse). This corresponds to the SCA approach discussed earlier. Typically, only voxels surpassing an arbitrary or statistically determined threshold on the correlation coefficient are included in the presented network. Note that such a network is derived from information on connectivity of whole brain to one seed region, not of connectivity of each area to each.

This method can be altered simply by choosing a different connectivity measure than correlation coefficient (e.g. mutual information, spectral coherence, or wavelet decomposition parameters correlation). The potential advantage of these more complex, potentially nonlinear methods for bivariate functional connectivity indices is critically assessed in chapter 5 of this thesis.

As a typical result of this method is a set of locations connected to a given seed region, the need to introduce a seed region based on prior knowledge becomes an inevitable limitation of this method. On the other hand, the SCA approach offers a straightforward interpretation in terms of statistical properties of the localised signals, which is not always the case for multivariate connectivity methods.

Multivariate methods

Under the umbrella term of multivariate methods we cover two main groups of analysis techniques. The first group includes the (typically linear) unmixing methods like ICA and Principal Component Analysis (PCA), while the second concerns various clustering methods.

Matrix unmixing methods (PCA and ICA)

In general, both PCA and ICA attempt to explain the data matrix X (each column representing a time series of one voxel) as a linear transformation of a matrix of source (latent) variables Y by a mixing matrix A :

$$X = AY. \tag{3.1}$$

If t denotes the number of acquired time points, v number of voxels, then X is a t -by- v matrix, A is a t -by- t matrix and Y is a t -by- v matrix. Clearly, such decomposition is generally not unique. In PCA, the constraint is that the first

column of matrix A shall explain most of variability of matrix X , and each successive most of the residual variability, given the condition of orthogonality to all the preceding vectors. This leads to decomposition such that the matrix A consists of the eigenvectors of matrix X in decreasing order of magnitude. Apart from decomposing the data to spatial maps (rows of Y) with assigned orthogonal time series (columns of A), the algorithm has the advantage of helping with data-reduction, since the first few principal components contain most of data variability, while the last ones can be discarded as usually containing only minor noise variance.

The ICA method decomposes the matrix Y based on different constraints. In particular, it poses the assumption of independence of the sources (either in spatial or temporal domain, for fMRI data usually the former is used). Note that this assumption is stronger than that of orthogonality in PCA, (which is equivalent to purely linear independence in case of demeaned data).

So far we have assumed that the number of estimated sources is the same as the number of time points we acquired. It is clear that this assumption does not generally have to hold, and might also be difficult to assess. In fact, the necessary decision on the number of sources to recover is one of the most controversial points of using ICA. Either it has to be based on a model, theory or intuition of the analyst, or a method to automatically estimate the number of sources contained in the data has to be implemented. Another problem (of both ICA and PCA) is that these models do not include explicit error terms, although some measurement error or noise can be expected to be always contained in the measured data.

We shall discuss here one solution to these burning questions concerning the use of ICA for fMRI data, which was proposed by Beckmann et al. [10]. Their Probabilistic Independent Component Analysis (PICA), implemented as the MELODIC routine in the FSL software package, uses estimation of source number by comparing the eigenvalues of the data covariance matrix with the distribution of eigenvalues of a random white noise covariance matrix. While this constitutes a meaningful criterion for data dimension reduction, we shall bear in mind that the solution proposed is still somewhat arbitrary. Moreover, it has been proposed that principally there *“can be no single, ‘best’ dimensionality or model order for the underlying neurophysiology of multiple distributed systems”*[33]. Importantly, the statistical model order selection may lead to over-splitting of components in data from longer acquisition with more timepoints,

which we have already observed in several cases in our experimental data having 430 timepoints in comparison to commonly published data having around 200 timepoints and resulting to less components. On the other hand, the dependence of typical number of detected components on the number of timepoints can be utilised in group studies using the concatenation technique such as in [98] to enable detection of splitting of the major resting state network into subnetworks. The advantage of ICA over the seed-based methods in detection of networks was documented in some studies [10, 114].

Clustering methods

Although we have not used clustering methods directly in our analysis, we include them in this review as they are starting to be used in resting state literature alongside ICA. We will give just brief overview of these methods, followed by example results from literature in the latter subsection.

Clustering methods generally attempt to group objects (or cases, in our context voxels or regions of interest) into several sets/clusters based on their similarity or dissimilarity. Clustering methods can be roughly divided into two approaches: *hierarchical clustering* and *partitional clustering*. [169]

In hierarchical clustering, typical hierarchical tree architecture is created by successive processing of previously established clusters, while partitional clustering algorithms determine all clusters at once, although potentially by an iterative process. Hierarchical algorithms can be agglomerative (“bottom-up”) or divisive (“top-down”). Agglomerative algorithms begin with each element as a separate cluster and merge them into successively larger clusters. Divisive algorithms begin with the whole set and proceed to divide it into successively smaller clusters. The simplest algorithm for hierarchical clustering starts with each case being considered as a separate cluster. Then the criterion on the uniqueness is gradually relaxed, and any two cases which are nearer to each other than a given threshold are merged (amalgamated) together into one higher-order cluster. The algorithm proceeds until all cases are merged in one big cluster. The result can be easily visualized by a hierarchical tree. The resulting tree is dependent on the particular distance measure used and the way of defining the distance of two clusters. For the distance, typically a single linkage (nearest neighbour), complete linkage (furthest neighbour) or some sort of average linkage method is used.

A typical example of a partitional clustering method is k-means clustering. Before running the algorithm, the analyst has to decide on the number k of clusters to be detected. The algorithm then tries in an iterative manner to assign each of the cases to one of k groups to maximize the ratio of between groups variance to within groups variance. Note that there are methods to estimate the most suitable number of clusters to be extracted directly from the data such as Bayesian Information Criterion (BIC). Recently, another algorithm for partitional clustering, so-called Normalised Cut Group Clustering, has been successfully applied to resting state data, detecting RSNs at group level [174].

Combined methods

In real-world analysis, several concepts are commonly combined in the analysis of spatiotemporal structure of resting state data. In the following we describe two examples of such an approach.

A nice example of the combined methodological approach to resting state dataset analysis is found in work of Thirion et al. [173]. In their multi-step approach, they:

- Spatially under-sample the data using intra-session parcellation algorithm, ending up with around 1000 parcels representing spatially and temporally homogeneous units
- Use pair-wise spectral coherence of parcel timecourses as a measure of connectivity between parcels
- Transform the pair-wise coherence matrix into a distance map using several different methods
- Run PCA on the distance map in order to reduce the dimensionality of the space by using only the first 3 PCA components
- Determine the number of clusters to be derived by Bayesian Information Criterion
- Run clustering algorithms (General Mixture Modelling) to obtain final clusters
- Use prior knowledge to choose the anatomically most plausible clusters for display

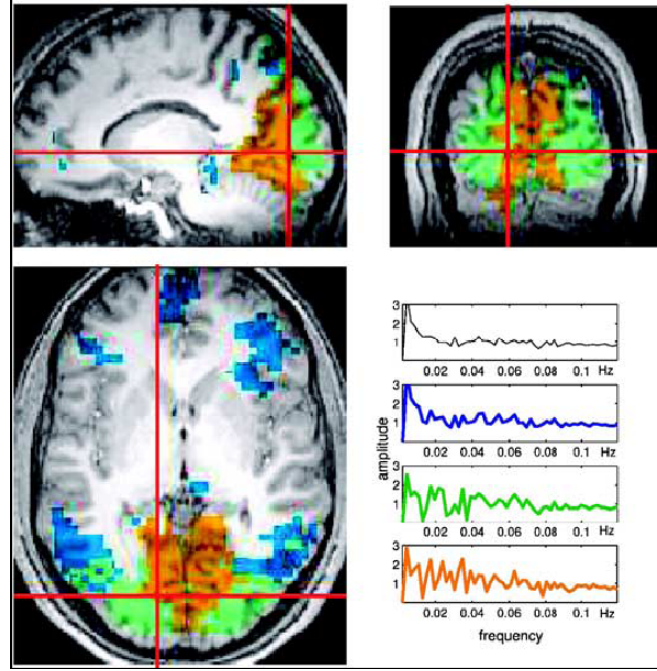


Figure 3.4: Figure shows the spatial and frequency distribution of three of the detected clusters by the complex method of Thirion et al. The first spectra (in black) corresponds to the average signal from the remaining unclassified voxels. Adopted from [173].

The results of this method are shown in Figure 3.4.

Another example of a complex approach appears in the paper by Achard et al. [1]. The steps of their analysis are as follows:

- Anatomical template based parcellation into 90 brain regions of interest followed extraction of the average time-courses for these 90 regions
- Wavelet analysis on six different frequency scales (defined by typical frequencies represented by the fitted wavelets)
- For each scale, a correlation matrix of the wavelet coefficients was computed, representing further the connectivity matrix
- Each connectivity matrix was thresholded controlling for False Discovery Rate at 5% level
- The thresholded connectivity matrix was binarized to serve as neighbourhood matrix for graph analysis using various indices, namely computing characteristic path length, clustering coefficient and degree

- An index of small-worldness was derived by comparison with typical characteristic indices of comparable random graphs

While the combined methods inevitably differ in many aspects, there are some common features:

- Spatial down-sampling of the input data
- Use of sophisticated connectivity/distance measures, possibly sensitive to coherence in a specific frequency band
- Need to reduce the dimensionality of the final results before a reasonable interpretation or visualization is feasible

Overall, these combined approaches have not been widely adopted, probably due to the complexity of the analysis and their dependence on many decisions regarding the settings of the analysis. On the other hand, if well tuned to a particular dataset, combined methods might gain an advantage over the general purpose approaches such as seed-based correlation or ICA.

3.3.2 Effective connectivity analysis

Effective connectivity methods have neither been widely used in this thesis, nor extensively applied to resting state fMRI data by others. Therefore in the following we provide only a brief overview of the approaches with relevant references for the sake of completeness.

As introduced earlier, there is an important distinction between functional connectivity, usually considered as ‘temporal correlations between remote neurophysiological events’ and *effective connectivity* characterised as ‘the influence one neural system exerts over another’ [56]. The study of effective connectivity seeks to recover a potentially asymmetric matrix of directed interactions between localised neural units. In theory, these neural units would correspond to neurons or neural populations, while in practice the connectivity between voxels or regions of interest is studied.

The effective connectivity methods differ in the form of the interaction model they assume. In its basic form, structural equation modelling (SEM) assumes linear instantaneous relations and independence of the temporal samples. When model of the dynamics is included, we deal with multivariate autoregressive modelling (MAR). Such a model allows implementation of testing of causality in the

sense of Granger [70], as the temporal precedence of cause and effect can be assessed. A richer model of causality-based effective connectivity is the Dynamic Causal Modelling (DCM) [60], which assumes a bilinear dynamic model on the level of neural population activity and includes an explicit forward model from the neuronal activity to the measured data. This model includes the haemodynamic response function in the case of fMRI data. In theory, a fully nonlinear nonparametric approach to effective connectivity analysis based on conditional mutual information or *transfer entropy* [157] could also be used, although the typically low number of datapoints in fMRI datasets poses a limitation on the robustness of estimation of information-theory based functionals.

In general, effective connectivity methods are less suitable for large-scale exploratory analysis and are therefore applied to signals from several well-defined areas of interest. Also, as the timing is crucial for the dynamic effective connectivity models, the spatial heterogeneity of the haemodynamic response function and the relatively low sampling rate generate serious difficulties to effective connectivity estimation and interpretation. In fact, there is an ongoing heated debate regarding the advantages and suitability of particular effective connectivity methods - see e.g. [57, 153].

3.4 Alterations in spontaneous brain activity

The phenomenon of spontaneous brain activity as observed with resting state fMRI did not remain unnoticed by wider community of researchers and clinicians, who raised the question of stability or modification of the low-frequency fluctuations and connectivity patterns represented by RSNs by factors such as mental state or disease. In this section, we summarize the findings of alterations of spontaneous brain activity. Special attention is paid to those related to changes of state of consciousness, as these are particularly pertinent to the research presented in chapters 4 and 6 of this thesis.

3.4.1 Alterations related to vigilance

One of the fundamental goals of modern neuroscience is to shed more light on the neural basis of consciousness. To investigate altered states of consciousness, studies of stimulus-free or “resting-state” fMRI are becoming increasingly popular. To determine mental-state-dependent alterations of the functional architecture of the resting brain, the two main properties of BOLD signal fluctuations

that were introduced earlier are typically assessed: FC capturing the extent and strength of spatio-temporal synchronisation and the strength of the underlying LFF.

FC fMRI studies during altered states of consciousness have so far generated somewhat controversial observations. For instance, while [99] reported an increase of FC in several areas including the motor network after induction of midazolam sedation, another group [142] reported decrease of FC in the motor areas in sedation induced by sevoflurane.

The DMN has previously been hypothesised to serve functions that may be related to conscious mental activity. Thus, the pattern of FC changes in higher cognitive networks and the DMN are currently being actively researched to unveil the neural basis of consciousness. So far, there was a decrease in DMN FC observed under midazolam sedation [72] and in brain-damaged patients [18]; inconclusive results were reported in light sevoflurane sedation [119]. Fitting in the general picture of partial reduction of DMN FC in suspended consciousness, DMN FC was shown to be preserved but reduced during deep sleep [86, 155], with predominant reduction in frontal loading.

In contrast to complex FC changes, reduced vigilance was found to be consistently associated with increases in LFF amplitude as studied during midazolam-induced sedation [99], light [62, 87] and deep sleep [106]. This is surprising in view of the opposite effect of low frequency vasomotor activity during anaesthesia [89, 88]. Vasomotor activity here refers to oscillatory vascular rhythms that share many properties with BOLD LFF such as frequency characteristics and responsiveness to manipulations of vascular tone. As said, in contrast to BOLD LFF in humans, the animal studies cited above showed a clear and significant dose-dependent reduction of vasomotion during anaesthesia. Intriguingly, a recent rodent fMRI study largely replicated these effects for BOLD LFF demonstrating reduction of both amplitude and coherence strength of LFF during deepening of isoflurane anaesthesia [112].

To reconcile this controversy about LFF changes in altered vigilance state, one might speculate on true species-specific effects. On the other hand, one may look for effects of experimental factors differently controlled during human and animal studies. BOLD LFF is known to be potentially confounded by a number of non-neural factors such as systemic cardio-vascular physiological variables, and artifacts such as motion. It has been noted that motion may affect LFF [164, 151], but no systematic study on such effects exists. As sedation and light sleep

may increase the propensity for head motion by reducing the subject’s conscious control of head position, we hence hypothesised that LFF increases observed in human sedation studies may be mediated by sedation-induced increased head motion. A study testing this hypothesis is described in Chapter 4. We also include some further analysis of the potential mediation effect of motion on FC in similar context in chapter 7.

3.4.2 Alterations related to disease

The properties of resting state fMRI signal, in particular in the DMN, have been extensively investigated as potential biomarkers of mental diseases or other conditions related to structural and functional brain changes. The spectrum of conditions for which changes of DMN properties have been reported is wide, including among others Alzheimer’s disease, Schizophrenia, Attention Deficit/Hyperactivity Disorder, Depression, Anxiety, Epilepsy, Autism Spectrum Disorder, Multiple Sclerosis and Parkinson’s disease. In most cases, decreases of FC and/or fluctuation power within the DMN are reported. For reviews and meta-analysis we refer the reader to [96, 6, 23]. Here we give only an example of two studies carried out by the group of Michael Greicius.

The first is a study on Alzheimer’s disease [73], which has revealed different spatial distribution (namely hippocampal co-activation) and overall higher connectivity in healthy elderly compared to incipient Alzheimer’s disease. They also proposed similarity to healthy brain template spatial distribution (derived from earlier studies) as a possible disease marker. This achieved 85% sensitivity and 77% specificity. These discrimination characteristics are quite promising, given the fact that the template was not optimised and this was a first study on the topic.

The later depression study [71] states that depressed subjects showed increased default mode connectivity in the subgenual cingulate, the thalamus, the orbitofrontal cortex and the precuneus. On top of that, cingulate connectivity correlated with length of depression. Nevertheless, in this study, goodness of fit to template was not a good biomarker of depression.

When interpreting the SBA and the alterations of SBA features with disease or brain state, the question of reliability of estimation of the SBA measures naturally comes up. Just recently, several studies focused on this topic [160, 126, 183]. These findings suggest that in general, the resting state network

spatial patterns show moderate to high reliability. This supports the potential for construction of disease biomarkers. Nevertheless, while some results of the clinical studies of alterations of spontaneous brain activity are promising, the sensitivity of measurement and analysis methods and our understanding of the nature of the phenomena is not sufficient yet for effective clinical applications.

3.5 Neurophysiological underpinnings

In view of the amount of research invested in the properties and potential clinical applications of spontaneous brain activity, it may be surprising how little is known about the foundations of these phenomena. The nature of the synchronised spontaneous LFF underlying resting state fMRI is still only poorly understood. Nevertheless, there is a growing number of studies researching the electrophysiological underpinning of certain aspects of these resting brain activity features.

In particular, band-limited local field potentials, as the best electrophysiological correlate of BOLD signal [113], were found to show LFF in anaesthetised monkeys [110].

In humans, the link between EEG band-limited power in particular bands and the fMRI signal has been investigated since the main technical problems with simultaneous EEG/fMRI recordings have been solved. Initial whole brain correlation studies between band limited EEG power and BOLD signal amplitudes yielded variable results with some studies localizing spectral EEG/BOLD correlations to the default mode [108]. More recently, negative correlations between frontal theta EEG independent component band-limited power and BOLD signal were reported in areas of DMN [156]. A more comprehensive picture was offered by a study focusing on resting brain networks and their EEG signature [118]. This study reported correlations of EEG power spectra with BOLD activity in specific networks. Correlations with DMN time course as derived by ICA were found to be generally positive, particularly for beta and alpha rhythms. Another recent study [90] conducted inverse analysis and found nodes of BOLD activity within DMN correlated to EEG global field synchronisation.

Generally, the previous EEG/fMRI studies focused on correlating the instantaneous EEG-band-limited power (or synchronisation) with BOLD signal intensity. Band-limited EEG power is however an indicator of local fast-scale synchronisation of a specific type of processing, rather than a mere measure of the intensity

of processing. Therefore we hypothesise that EEG spectral powers as an indicator of brain state may be more directly linked to some network feature such as FC. This link may in fact be stronger than the observed but variable link to metabolic intensity measurement of BOLD intensity, and allow an intuitive investigation of associations between network-specific FC and EEG power. Moreover, focusing on interrelations between EEG power and inter-regional FC of BOLD signal would enable the direct study of inter-individual and inter-session variability that may underlie altered network organisation in relation to personality traits, emotional states and importantly diseases. To the best of our knowledge no previous study investigated the interrelation between EEG power and synchronisation of fMRI activity during awake state directly. Therefore we set out to test the hypothesis that the strength of synchronization of BOLD activity within the DMN may be related with band-limited EEG power. This study is described in chapter 6.

Finally, a very promising approach for EEG/fMRI correlation was presented in two recent studies [21, 133]. These studies propose correlating the BOLD signal with time-courses derived from conventional EEG microstate analysis and show that particular fMRI RSNs have time series correlated to the instantaneous density of particular microstates.

3.6 What might spontaneous brain activity represent?

As reviewed in the previous section, the understanding of the nature of the spontaneous brain activity is still in its initial stages. While cognitive correlates have been found for some of the markers derived from spontaneous brain activity fluctuations such as in [120], it seems that unconstrained thought and other higher order cognitive processes are not at all sufficient for explaining this activity.

To this end, interesting results were provided by observation of BOLD fluctuations in non-human species, namely monkeys. Resting state networks have been proved to be present in anaesthetized monkeys even at anaesthetic levels known to induce profound loss of consciousness [176]. Furthermore, not only oculo-motor and somato-motor networks were detected, but an analogue of human default network has been revealed by regression with seed in the precuneus cortex. The spatial pattern of connected areas fairly well resembled that of human default network. These results support the position that spontaneous brain

activity, particularly in DMN, does not correspond merely to conscious mental activity such as self-referential thought or daydreaming.

So what may the spontaneous brain activity represent? This is so far an open question; nevertheless, many speculations have been suggested. For now, we may share with the reader some of the interpretations, offered in [149]. The first is that spontaneous brain activity represents spontaneous cognition e.g. stimulus independent thoughts. This may be the case for some of the fluctuations, but clearly cannot account for the low frequency fluctuations during anaesthesia or deep sleep. The other concept mentioned is that intrinsic activity facilitates maintaining of balance between excitatory and inhibitory inputs determining the responsiveness of neurons to correlated input. Finally, the most appealing interpretation is that: *“the intrinsic activity instantiates the maintenance of information for interpreting, responding to and even predicting environmental demands. In this regard, a useful conceptual framework from theoretical neuroscience posits that the brain operates as a Bayesian inference engine designed to generate predictions about the future”* [149].

Part II

Analysing Spontaneous Brain Activity

Motion as a confound in resting state studies

As already discussed in Section 2.2, resting state fMRI data suffers from the presence of structured noise (artifacts) due to various sources. In this chapter, we discuss this issue in more detail and present some analysis of the potential confounding effects of these artifacts on studies quantifying changes of LFF power. In particular, we focus on the potential role of subjects' head motion. The work presented in this chapter has been published in the Magnetic Resonance Materials in Physics, Biology and Medicine journal [79].

4.1 Nonneural sources of LFF

4.1.1 Overview

Apart from the neuronal sources of resting state BOLD signal and the inevitable thermal noise, several non-neuronal (or artifactual) sources are usually considered [11, 175] including instrumental drift, participant motion, and physiological sources such as cardiac and respiratory activity (and variability of their rate). While these non-neuronal sources are not specific to resting state fMRI studies and are present in active-task studies as well, in resting state studies they gain on importance as no explicit model for neuronal activations is available and therefore we miss this useful prior information for distinguishing the artifactual activations from true neuronal activations. The relative contribution of neural and various non-neural mechanisms to BOLD fMRI signal variations has been investigated in a range of conditions [11, 164]. In general, the need for suitable preprocessing was firmly established - we review the preprocessing options for

resting state fMRI in the following subsection.

4.1.2 Typical preprocessing procedures

Let us recall that basic fMRI preprocessing consists of removal of the first few 3D volumes to avoid initial signal instability, motion correction by means of rigid body transformation of each 3D volume to achieve the best match to a selected reference volume (typically the first, middle or last one of the acquisition) and potentially the so-called slice-timing correction, which is basically a temporal interpolation carried out to correct for difference in acquisition times of different volume slices. Quite often a so-called coregistration is also carried out by means of an affine transformation of the brain image to achieve a best match to a standardised template brain (typically the Montreal Neurological Institute (MNI) template brain [48] is used). This is done in order to allow group statistics comparison, as well as for the reporting of results in terms of standardized coordinates. Commonly, some spatial smoothing is also applied to increase the signal-to-noise ratio in exchange for spatial specificity. While the optimisation of all the above steps is a complex task in its own and subject to further research, the state-of-the-art implementation of the necessary algorithms is available and regularly updated in the main software packages - such as Statistical Parametric Mapping (SPM) [58] or FMRIB Software Library (FSL) [166].

Additional preprocessing steps are typically carried out in order to remove, at least partly, the remaining non-neural sources of temporal signal variation. As discussed earlier, in resting state fMRI, the signal of interest has typical low frequencies (about 0.01 – 0.1Hz). On the other hand, the cardiac and respiratory activity has considerably higher typical frequencies (~ 1 and ~ 0.2 Hz respectively). Therefore the use of low-pass filters with cut-off frequency about 0.1 Hz is widespread in fMRI preprocessing. Similarly, since slow scanner signal drifts have been reported, a high-pass filter at about 0.01Hz is commonly used.

Importantly, the band-pass filtering can be insufficient - with typical TR of 2 seconds for whole brain acquisition, the Nyquist frequency is about 0.25 Hz and the cardiac and potentially the respiratory signal can be critically undersampled, leading to aliasing of the artifact, potentially into the 0.01 – 0.1Hz frequency band of interest. While the work of Biswal [15] or Cordes [36] suggests that this aliasing does not play a major role, at least in the FC analysis of resting state data, approaches have been proposed to solve this problem by modelling the non-neural contributions utilising simultaneously acquired physiological data

[14, 41, 67]. Another approach lies in the use of ICA, which was reported to be able to separate the physiological artifacts from resting state networks based on their spatial pattern even under conditions of temporal undersampling [38].

An alternative approach to removing signal components due to physiological artifacts is by orthogonalising all the voxel or region of interest (ROI) timecourses with respect to the average signal taken from ROIs in white matter and cerebrospinal fluid. The rationale of this step lies in the assumption that as these tissues do not generate true neuronal signals, their signal fluctuation solely reflects the artifactual sources. The orthogonalisation or regressing out is a simple application of a linear regression - effectively removing the part of the signal Y that is correlated to the estimated artifactual signal X , as given by:

$$Y_{orth} = Y - (X(X^T X)^{-1} X^T)Y.$$

Here T in the superscript stands for matrix transpose and X , Y are considered column vectors of size n samples. If there are several artifactual signals $X_1 \dots X_n$ to be removed, they can be regressed out using the same formula by putting the signals into columns of a single matrix X . Note that this formula is directly applicable only in the case that the signals Y and X have zero mean - otherwise a constant has to be included in the regression model by adding a vector of ones to the X matrix.

A similar procedure is commonly used to decrease the extent of motion-related signal variation that was not removed by initial motion correction by spatial realignment of the images. Here, 6 motion parameter time series estimates are most commonly used as regressors, although the inclusion of the second order and lagged effects has also been proposed to correct for nonlinear and lagged effects of head-motion [61].

It is not uncommon that artifacts lead to widespread positive correlations, affecting globally the whole brain image and impacting on reliability of FC analysis. The global character of these fluctuations motivates another method of data preprocessing, namely global signal correction - correction for fluctuations averaged across the entire brain. While at least three different approaches to this correction have been used (for a review see e.g. [52]), the most current reports typically utilise a method based on regression of the global signal using an orthogonalizing procedure similar to the one described above for removal of physiological artifacts and remaining motion-related signal variations. The use of global signal correction has recently attracted a heated debate in the resting

state neuroimaging field regarding its suitability. Most importantly, global signal correction was an important preprocessing step in original studies that reported anti-correlated networks [51, 53]. The interpretation of the FC results, and particularly the anti-correlations, after application of global signal correction, has been questioned [132]. The main argument here was that the global signal correction, purely due to the mathematical consequences of its construction, forces the distribution of voxel values of seed based correlations to be approximately centred around zero. In other words, independently of what the true correlations in the neuronal signals are, after applying the global signal correction we are bound to detect a comparable amount of positive and negative correlations. The global signal removal not only removes global artifacts, but also signals correlated to average brain neural sources, obscuring the interpretation of the results. In particular, it could be argued that the anti-correlations observed after global mean regression are not truly neuronal, but introduced by the correction [132]. While the reply of some of the strongest supporters of global signal correction [52] may not seem completely convincing, independent evidence brought by [27] has confirmed that at least some of the anti-correlations are genuine, by documenting these anti-correlations in the absence of global signal correction. This report advocated the use of regression of physiological signals or their image-based counterparts as a safer alternative to global signal correction.

Similarly, an image-based artifact removal procedure consisting in regressing out average signals from white matter and cerebrospinal fluid regions and motion parameters was proposed as “an effective method for dealing with global correlations that offers the significant advantage of reducing the risk of introducing ‘fictitious’ negative correlations in the data sets” [66].

4.2 Is sedation induced increase in low-frequency fluctuation mediated by increased motion?

Having reviewed the typical fMRI preprocessing methods, the question arises, whether these are fully sufficient for any type of subsequent data analysis. In the following, we study the potential role of unbalanced artifacts in between-condition comparisons.

4.2.1 Introduction

As mentioned earlier in Section 3.4, there are an increasing number of studies that focus on changes in the crucial resting state signal properties - either due to specific disease conditions or due to other manipulations of brain state. Assessing the induced changes in the main resting state features, such as low frequency fluctuation power and functional connectivity, should help to extend our understanding of the neurobiological underpinnings of consciousness and also the spontaneous brain activity itself.

On the other hand, resting state fMRI signal is prone to artifacts which are difficult if not impossible to remove completely. Therefore, a reasonably imaginable scenario for the observed changes in resting state measures in altered brain states is that these constitute a result of an increase in some of the artifact rather than in the true neuronal effect. Apart from the extreme scenario of no neuronal effect and observed change fully attributable to change in the artifacts, the mixed scenario of an artifact change partially increasing the estimate of a real change, or even masking a change in the opposite direction seems even more likely. To complicate the situation further, a spatial selectivity of the artifacts or signal changes might lead to spatially heterogeneous effects of the artifacts, masking out the changes in some networks while increasing the changes in other networks.

To explore this scenario more specifically, we hypothesised that LFF increases observed in human sedation studies may be mediated by sedation induced increase in head motion. To test this hypothesis, we studied LFF in sensory and cognitive networks in healthy volunteers during baseline and midazolam induced conscious sedation. Mediation analysis consisted of testing a direct sedation effect and indirect motion effect with motion indexed as mean relative displacement. In addition, motion effects on LFF were directly quantified during baseline state (i.e. before sedative administration), representing the standard experimental resting state fMRI setting.

4.2.2 Material and methods

fMRI data from healthy volunteers were included that had been acquired in the context of an unrelated study on the cerebral effects of midazolam. MRI scanning was performed on a Phillips Achieva (Philips, Eindhoven, NL) at 3T. The study was approved by the Nottingham University Medical School ethics

Table 4.1: Ramsay sedation scale ([150])

Awake levels are classified as:	
Level 1	patient anxious and agitated or restless or both
Level 2	patient co-operative
Level 3	patient responds to commands only
Asleep levels are dependent on the patient's response to a light glabellar tap or loud auditory stimulus:	
Level 4	a brisk response
Level 5	a sluggish response
Level 6	no response

committee and all subjects gave written informed consent. 20 healthy volunteers (18 males, 2 females, age range 18-35) were scanned during resting wakefulness (baseline). Subjects were instructed to lie still with eyes closed and not to fall asleep. 15 of these volunteers underwent a second session on the same day after induction of light sedation by intravenously injected midazolam ('Dormicum', a benzodiazepine derivative, 0.05mg/kg). Volunteers were rescanned within 5 minutes after they stabilized at Ramsay level 3 (see Table 4.1 for description of the Ramsay sedation scale [150]) that was assessed (by the attending anaesthesiologist) with the subject on the scanner table outside the magnet with the head coil removed.

430 volumes of standard functional single-shot echo-planar images (EPI; TR = 2100ms, TE = 60ms, flip angle 90°, 64x64x35 matrix, resolution $3.25 \times 3.25 \times 3$ mm) were acquired for each scan. Standard pre-processing steps were applied, using FSL4.0 [166] software package: brain extraction, motion correction and high-pass 0.01 Hz frequency filtering. For each session, registration matrices to MNI template were computed and their inverses used to register MNI-space based ROIs to individual acquisitions. The ROIs were chosen to represent the main nodes of 5 functional networks: 4 ROIs for the DMN, 4 ROIs for the DAN and two for Auditory, Motor and Visual Network.

The following procedure for deriving the DMN and DAN ROIs was used: firstly, for each network a binary network template consisting of spherical regions (15 mm diameter) centred at significant network peaks was created. The peak locations were adapted from SI Table 3 of [118]. Secondly, individual ICA decompositions of MNI-registered fMRI data were obtained using MELODIC (FSL4.0) [10], utilising built-in automated estimation of number of components. Subsequently, a template-matching procedure proposed in [73] was run on the decompositions

ROI	x	y	z
DMN: PCC	0	-50	18
DMN: MPFC	0	52	28
DMN: IPL left	-50	-60	24
DMN: IPL right	50	-60	24
DAN: PL left	-40	-64	44
DAN: PL right	40	-64	44
DAN: DLPFC left	-42	26	36
DAN: DLPFC right	42	26	36
Motor: left	-56	-2	38
Motor: right	56	-2	38
Visual: left	-8	-76	12
Visual: right	8	-76	12
Auditory: left	-42	-24	12
Auditory: right	42	-24	12

Table 4.2: MNI coordinates of template 15 mm sphere ROI centres. PCC: precuneus cortex, MPFC: medial prefrontal cortex, IPL: inferior parietal lobule, PL: parietal lobule, DLPFC: dorsolateral prefrontal cortex

to detect best-matching individual maps for each binary network template. As a last step, for each network, all the individual best-matching components were averaged and the resulting group map thresholded at $Z > 2$ and binarized. The resulting binary network masks were then split into several ROIs. In particular, 4 ROIs were created for DMN located in precuneus/posterior cingulate cortex, medial prefrontal cortex and left and right inferior parietal cortex and 4 ROIs for DAN network (bilateral dorsolateral prefrontal ROIs and parietal ROIs). These 8 ROIs were complemented by 6 ROIs (3 left-right pairs) from Auditory, Motor and Visual Network. These ROIs were derived with analogous procedure with the difference that the centres of the template spheres were located at hemispherical peaks of anatomical maps in probabilistic Harvard-Oxford Atlas included in FSL 4.0: the initial template for the Auditory Network was based on Heschl's gyrus map, for the Motor Network on precentral gyrus map and for the Visual Network on intracalcarine sulcus map. All ROIs were checked for spatial overlap between them and the overlapping areas (although not constituting substantial parts of any of the ROIs) were excluded from the ROIs. The centre locations of the sphere templates are summarised in the Table 4.2, while Figure 4.1 shows the derived ROIs.

To analyse the potential impact of motion on resting state fMRI analysis, we followed standard procedures for data preprocessing. Average time courses for each ROI of each subject were extracted and orthogonalized with respect to mo-

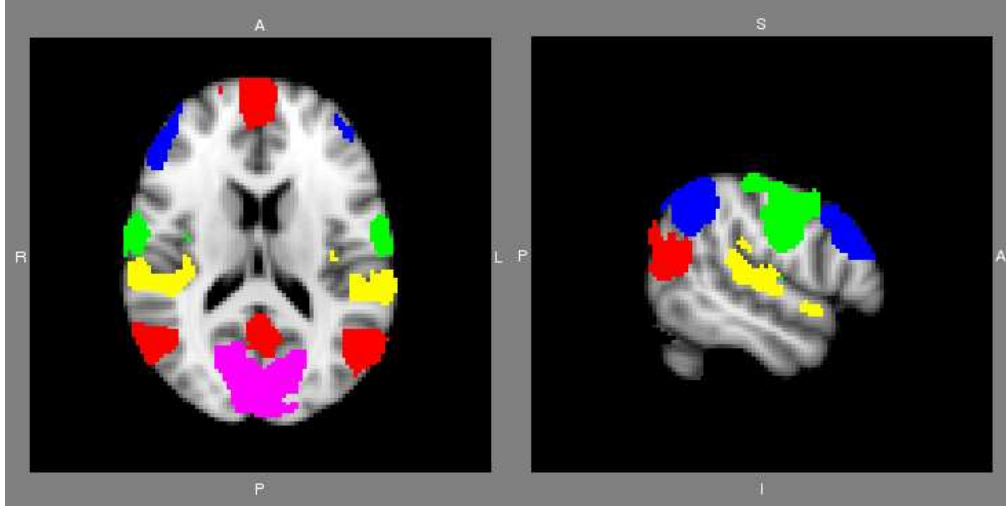


Figure 4.1: ROIs derived for five RSNs: auditory (yellow), DMN (red), DAN (blue), motor (green), visual (pink). Each network is represented by 2 or 4 ROIs. On the left is an axial slice ($z=16$), on the right a sagittal slice ($x=18$)

tion parameters and also with respect to average white matter and cerebrospinal fluid (CSF) time series to correct for nonspecific global signal fluctuations. The motion parameters were obtained from the standard 6 degrees of freedom spatial realignment, i.e. rigid body transformation motion correction procedure, while the white matter and CSF time courses were derived from thresholded white matter and ventricular ROIs from the Harvard-Oxford Atlas included in FSL 4.0. Finally, a low-pass 0.1 Hz second-order Butterworth filter was applied. Combination of FSL tools and in-house MATLAB scripts was used for the analysis.

The procedure for quantifying the LFF, implemented in an in-house MATLAB script was as follows: First, each average ROI time series was divided by its mean for normalisation. Subsequently, the absolute values of its FFT coefficients were squared and summed over the frequency bins within the 0.01-0.1 Hz range. The sum represents a measure of LFF power for a given ROI of a given subject. To derive a network-specific LFF power estimate, LFF measures were averaged across ROIs constituting a given network.

To further investigate the effect of global mean correction demeaning as an additional pre-processing step, we repeated the analysis with an extra orthogonalization step with respect to each session's global mean time series. We refer to the resulting LFF estimates in further text as 'global mean corrected'. The head-motion amount was measured by Mean Relative Displacement (MRD) [91], an

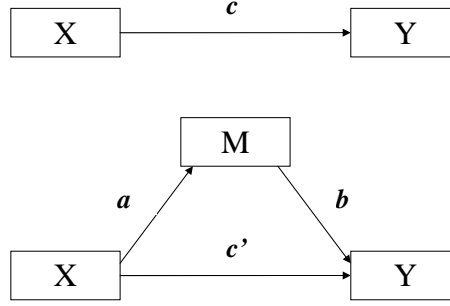


Figure 4.2: Illustration of the mediation model. Top: direct effect - independent variable X (IV) affects directly dependent variable Y (DV). Bottom: mediated effect: X affects Y both with direct effect c' and through indirect effect mediated by mediator M ; the strength of the mediated effect is equal to ab . Modified from [115]

index implemented in FSL4.0, making use of head-motion correction parameters from the registration algorithm.

Testing mediation

According to Baron & Kenny [9], testing the full mediation consists of four steps outlined below (so-called *method of causal steps* - Figure 4.2 illustrates the model considered):

1. establishing existence of total effect, i.e. correlation between the IV and DV, (effect c)
2. establishing existence of effect of IV on the mediator (effect a)
3. establishing existence of effect of mediator on the DV (effect b)
4. establishing non-existence of direct effect of the DV on the IV, if the mediator is included in the model (effect c')

More formally, simple mediation is assessed through use of a linear model described below (adopted from [115]):

$$Y = i_1 + cX + e_1 \quad (4.1)$$

$$Y = i_2 + c'X + bM + e_2 \quad (4.2)$$

$$M = i_3 + aX + e_3, \quad (4.3)$$

where i_1 , i_2 and i_3 are intercepts, Y is the dependent variable, X is the independent variable, M is the mediator, c is the coefficient relating the independent variable and the dependent variable, c' is the coefficient relating the independent variable to the dependent variable adjusted for the mediator, b is the coefficient relating the mediator to the dependent variable adjusted for the independent variable, a is the coefficient relating the independent variable to the mediator and e_1 , e_2 and e_3 are residuals.

We structure the investigation as follows: first, we test the total effect (c) of sedation of LFF, in each of the five networks - this corresponds to the first causal step of Baron & Kenny. Subsequently we run a joint test of the second and third causal step by means of a bootstrapping procedure for mediation analysis described in [143] as implemented in SPSS script INDIRECT [144]. In brief, this test consists in comparing the value of the product of the coefficients a, b as defined above against the hypothesis $ab = 0$. This is in line with the common agreement that joint tests of a and b in mediation analysis are more highly recommended - for extensive comparison of mediation tests see e.g. [116]. For completeness, apart from the main joint test results, we include a separate account of effect of sedation on head-motion (a) and head-motion on LFF (b). Finally, we report the test of the direct effect of sedation on LFF (c'), when the indirect effect through motion is included in the model. Each of the individual coefficients a, b, c, c' was tested in the standard parametric framework of linear regression model, i.e. by a two-tailed t-test after dividing it by its standard error. As 9 of the 11 variables investigated have shown significantly non-normal distributions (Shapiro-Wilk test of normality $p < 0.05$), we have log-transformed the variables, including the motion estimate, before running the mediation analysis. The variables were also standardised to have zero mean and unit variance.

Robustness

To check the robustness of the mediation results with respect to the motion correction preprocessing step, we repeated mediation analysis on data with less and no motion preprocessing, firstly data without the motion orthogonalisation step and secondly data without any motion correction at all.

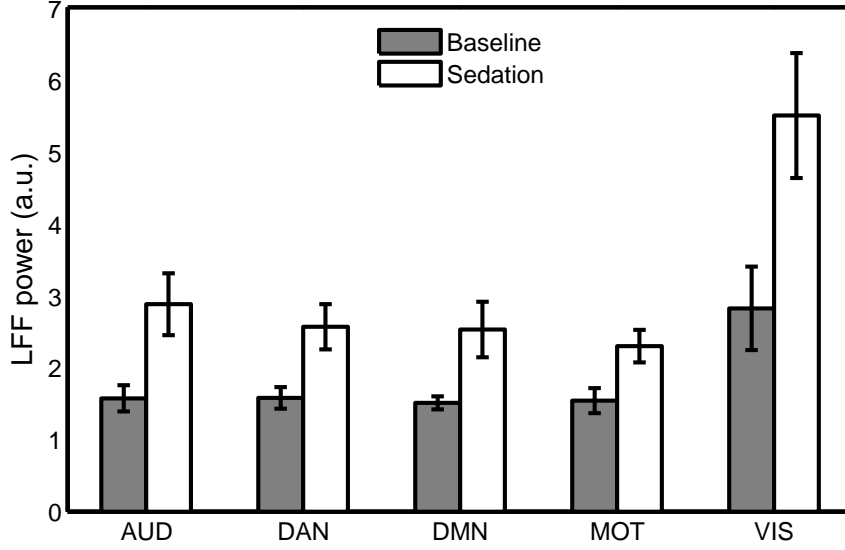


Figure 4.3: Effect of sedation on LFF power. The mean and standard deviation per network per condition is represented by full bar height and error bar. The effect proved significant ($p < 0.05$) in all five networks analysed: auditory (AUD), dorsal attentional (DAN), default-mode (DMN), motor (MOT), visual (VIS).

4.2.3 Results

Pharmacological sedation increases in LFF power

Sedation was found to increase LFF power (paired t-test, $p < 0.05$) of resting state BOLD time series in all sensori-motor and cognitive networks (visual, motor, auditory, DMN, DAN) (Figure 4.3).

Is effect of sedation on LFF mediated by motion?

For standard preprocessed data, we showed a significant indirect effect mediated by motion for all 5 networks considered ($p < 0.05$, corresponding to zero not included in the bias-corrected accelerated bootstrap 95 % confidence interval for ab). See a scatterplot in Figure 4.4 exemplifying the relation of the variables considered for visualisation of the mediation effect. To give a more detailed analysis, we split down the intervening effect into the two steps. Firstly, the effect a of sedation on motion was quite profound. The mean relative displacement in control condition was 0.07 ± 0.03 mm (mean \pm std) and in sedation 0.21 ± 0.10 mm, which constituted a statistically significant difference between the conditions ($p < 0.001$). Secondly, motion had a significant direct effect b on

the LFF power (t-test, $p < 0.05$) in all five networks.

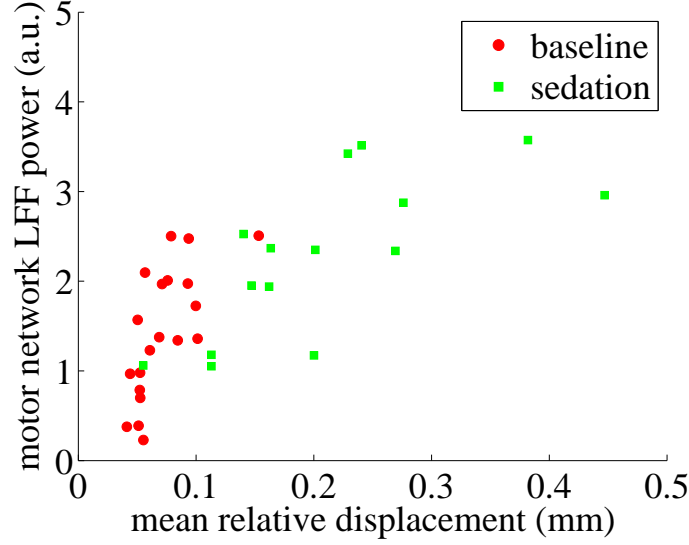


Figure 4.4: Visualisation of the mediation effect. The increase in LFF power (y-axis) between baseline (red circle) and sedation (green square) is significantly mediated by the sedation related increase in head-motion (x-axis), which itself elevates the LFF power even within baseline condition. This effect was found in all five investigated networks.

Absence of a direct effect of sedation on LFF

When the model included the mediating effect of motion, there was no significant remaining direct effect c' of sedation on LFF for any of the networks studied.

Results with additional global mean correction

Including additional global mean correction did not drastically change the results. In the first step, a similar but weakened total effect was noted: LFF power increased significantly in three (auditory, visual and motor) of the 5 networks with a trend for DMN as well ($p < 0.1$). In the next step, the indirect effect mediated by motion proved significant for the auditory and motor networks, with a trend in the visual network (95% bias-corrected accelerated confidence interval: -0.004 to 0.645). The indirect effect has shown insignificant in the two cognitive networks. Finally, the direct effect of sedation has not proven significant for any of the networks. The overall results are summarised in the Table 4.3.

Standard pre-processing	Sedation effect on motion (a)	Motion effect on LFF (b)	Direct sedation effect on LFF (c')	Total sedation effect on LFF (c)	Indirect effect mediated by motion (ab)
Auditory	0.73*	0.81*	-0.11	0.49*	0.60*
DMN	0.73*	0.50*	0.09	0.46*	0.37*
DAN	0.73*	0.54*	0.01	0.41*	0.40*
Motor	0.73*	0.96*	-0.31	0.40*	0.70*
Visual	0.73*	0.62*	0.06	0.51*	0.45*
Additional global mean correction:	(a)	(b)	(c')	(c)	(ab)
Auditory	0.73*	0.51*	0.23	0.60*	0.37*
DMN	0.73*	0.00	0.34	0.34*	0.00
DAN	0.73*	-0.08	0.28	0.22	-0.06
Motor	0.73*	0.61*	-0.04	0.41*	0.45*
Visual	0.73*	0.42?	0.21	0.52*	0.31?

Table 4.3: Summary of mediation analysis results. Columns show standardized regression coefficients in the mediation model. Significant effects are denoted by an asterisk (*, $p < 0.05$), trends close to significance by a questionmark (? , $p < 0.1$). For detailed explanation of the effects a, b, c, c' refer to the Materials and Methods section.

Robustness

To investigate whether the observed effects could have been introduced by inadequacies of the motion correction procedure, we repeated mediation analysis on motion corrected (realigned) data without motion orthogonalisation and raw data without any motion correction. In both cases, the estimated ab paths were similar or moderately stronger than those for the standard preprocessing and the hypothesis tests gave almost the same results.

Further, in a follow-up analysis we also included subject coding dummy variables as covariates in the mediation model to account for between-subject LFF baseline differences, which could potentially increase model sensitivity by removing between-subject variance. To put this into context, in this enriched model the tests of the coefficients c' and b correspond to the main and the confound effects in a repeated measures ANCOVA with varying covariate (ie. the motion amount). Nevertheless, inclusion of subject-specific baselines left the results almost unaffected. There was no change, or mild increase, in the estimated strength of mediated effect ab with respect to the original mediation model. The formal hypothesis test result changed only for one of the variables (DMN).

Counterintuitively, this was from significance ($p < 0.05$) to a trend or marginal significance ($p < 0.1$). This might have happened due to decrease in power of the test caused by lower number of degrees of freedom in the model that includes subject-encoding covariates.

Relations within baseline condition

To further investigate the relation between motion under standard experimental condition after standard motion correction procedures and LFF, we separately analysed data from the 20 baseline sessions, i.e. sessions when no sedatives were administered. We still observed significant ($p < 0.05$) positive correlations between motion parameters and LFF for auditory ($r=0.531$), motor ($r=0.651$) and visual networks ($r=0.680$). This corresponded to up to 46 % of inter-subject variance in LFF being explained by motion within a standard control condition of fMRI after using a standard motion correction scheme and motion orthogonalization. Even after additional global mean correction, motion was significantly positively correlated with visual LFF power ($r=0.601$); LFF in auditory ($r=0.353$) and motor ($r=0.402$) networks showed a positive correlation trend, although not statistically significant. Of note, DMN and DAN showed none or weak negative correlations of LFF to motion. In order to preliminarily explore the effect of head-motion on FC estimates, we have performed simple correlation analysis of motion and FC on the baseline dataset. The mean correlation across all ROI pairs was $r_{mean} = 0.22$, standard deviation $r_{std} = 0.10$, with maximal correlation $r_{max} = 0.66$.

4.2.4 Discussion

We have observed that pharmacologically induced light sedation increases LFF in resting state or stimulus-free BOLD fMRI. The effect was noted in all tested sensori-motor and cognitive networks, which is in line with the previously reported wide-spread LFF increases in sedation [99] and sleep [62, 87, 106].

Notably, we could verify our hypothesis that this observed sedation effect on LFF was mediated by head motion. In addition, a significant direct and strong effect of motion on LFF was confirmed by correlation analysis also during standard experimental baseline conditions without sedation. The demonstrated mediation effect supports the notion that the sedation-related increases in LFF to a significant extent can be explained by sedation-related increases in head motion.

In other words, motion is a significant potential confound in resting state fMRI studies even when using high quality data and state-of-the art motion correction algorithms. Additional global mean correction was found to mitigate, but not fully resolve this motion effect with significant residual mediation effect for sensori-motor networks. The results were also robust with respect to two assessed preprocessing alternatives (no motion orthogonalization or no motion correction at all) as well as inclusion of subject-specific baseline values of regional LFF in the model, documenting that the effects were not introduced by the analysis methods. While detailed comparison of motion correction approaches is beyond the scope of this study, these results allow excluding the possibility that the observed effects were introduced by the motion correction itself.

In our mediation model, midazolam induced light sedation was not associated with LFF increase when the indirect effect of motion was considered. This cautions against the interpretation of previously reported LFF increases during sleep and sedation as of neural origin, since beyond standard motion correction, motion parameters were generally not explicitly controlled for or studied. These findings do not contradict the possibility that LFF changes may be observed in states of reduced or even lost consciousness, but suggest that previous reports on LFF increases may have resulted from condition-specific subjects' propensity to head motion. This is particularly relevant as there is no neural mechanism known to date that would easily explain increased LFF during reduced vigilance. Lack of genuine LFF increase during conscious sedation would furthermore be well in line with rodent studies showing a gradual decrease of BOLD and CBF LFF during deepening of anaesthesia [88, 112] deploying experimental conditions with rigid head restraint.

Nevertheless we have to bear in mind that correct interpretation of a mediation test depends on the accuracy of the underlying causal model. At least two alternative explanations lend themselves to consideration. Firstly, motion might not actually be (the only) mediator of the LFF increase, but instead might serve as a marker for some other mediating physiological variable related to increased LFF, which has not been directly modelled, e.g. RVT (Respiratory Volume per Time) variability. Notably, even in such a case, it appears that motion estimates will give us important information regarding the level of non-neuronal LFF in the data where direct measurement of the underlying physiological variable may not be available. Secondly, one might argue that the explanatory power of motion estimates beyond sedation is still linked to neuronal LFF. In other words,

that motion may be a very sensitive measure of the underlying brain state - so sensitive, that it outperforms sedation manipulation. While this might seem counter-intuitive, we cannot disprove it based on our measurements. Also, our study does not allow us to speculate as to how motion may directly or indirectly affect LFF. Sophisticated experimental studies would be necessary to elucidate the exact mechanism generating the observed motion-mediated LFF increase.

Beyond sedation induced effects of motion, we found a strong correlation between head motion and LFF power also during baseline sessions reflecting standard experimental fMRI conditions in healthy young volunteers. This further warns against interpretation of LFF without accounting for the between subject or condition variation of motion. Even if there is no systematic bias between groups or conditions under investigation, the considerable amount of variance (46 %) of LFF power we could explain by mean relative displacement at the very least is prone to reduce statistical power. While this study focused on LFF power, due to this strong motion effect at baseline, we also explored motion effects on FC estimates. Our preliminary findings showed that at least in some pair-wise correlations, strong interdependencies can be observed between motion and FC metrics. This suggests that motion induced LFF power can translate into motion induced increases in FC estimates.

In practical terms, our findings call for extra caution in design, analysis and interpretation of FC MRI studies. This is especially pertinent for studies where the variable of interest (such as sedation) is related to bodily changes that may have an indirect effect on the studied ‘neural’ variable such as LFF power or FC. Removing maximally this noise variance with powerful post-processing techniques is important, but not necessarily sufficient, as the amount of residual noise is likely to still relate to the variable of interest. But also in studies where the variable of interest is not likely to alter head motion, statistical power of FC MRI analysis will be improved by extra experimental care to reduce head motion, more stringent rejection criteria and lastly explicit accounting for residual between subject/condition motion effects.

4.2.5 Conclusion

This study confirms previous reports that pharmacologically induced light sedation increases LFF in stimulus-free or resting state BOLD. The effect was noted in all tested sensori-motor and cognitive networks. Notably, we could verify our hypothesis that this observed sedation effect on LFF was mediated

by head motion. In addition, a significant direct and strong effect of motion on LFF in sensori-motor networks was confirmed by correlation analysis also during standard experimental baseline conditions without sedation. These effects were observed even after standard motion correction and after additional global demeaning, which calls for extra caution in FC study design, analysis and interpretation.

4.3 Summary

For the study of spontaneous brain activity and its changes in disease and altered brain state, one of the central methods is the assessment of strength of the low-frequency fluctuations. This has been reported to increase in conditions such as sedation or sleep. As there are other variables that may be affected by alteration of the brain state, we hypothesised that the observed fluctuation increase may be mediated by these confounding variables. A representative candidate of such a confounding variable is the amount of involuntary head-motion. We tested the hypothesis that head-motion mediates the observed fluctuation increases in sedation condition using a standard statistical test for mediation, exploring the inter-subject relations among amount of head motion, observed signal fluctuation and sedation condition. We have detected a mediating effect of motion for all five studied networks under standard data preprocessing, that persisted in at least two of the networks even after additional preprocessing steps. This provides an important warning for studies investigating spontaneous fluctuation changes.

Is linear correlation a suitable measure of functional connectivity?

Linear correlation is a generally accepted measure of FC in fMRI studies. Nevertheless one may ask whether linear correlation is indeed the right measure for FC. Does it describe well the degree of statistical dependence between the signals from different regions, or does it omit important information by focusing only on the linear component of the interdependence?

In this chapter, we propose a framework to investigate this question in a systematic way. Due to the threat of dimensionality explosion, we limit our investigation to instantaneous bivariate dependences. Nonetheless, we believe the results provide an important initial answer to a question which might be on the mind of many imaging neuroscientists. The core of this work has been already submitted for publication in the *NeuroImage* journal [81].

5.1 Introduction

As discussed in Chapter 3, probably the most widely used method of measuring FC between a pair of regions is computing a linear correlation of their activity time series. These activity time series can be derived from these regions by e.g. simple spatial averaging across all the voxels in the regions. Linear correlation is also widely used to obtain so-called correlation maps by correlating the seed voxel or seed region signal with signal from all the other voxels in the brain, potentially constrained to grey matter area. When we follow the distinction of FC methods

to bivariate and multivariate suggested in section 3.3, linear correlation has a dominant position among the former, while Independent Component Analysis is the most prominent example of the latter.

Indeed, from all possible bivariate measures of association, linear correlation is clearly a method of first choice. This reflects the commonly accepted assumption that the relationship between the fMRI time series can be suitably approximated by a multivariate Gaussian white noise process. Additionally, linear correlation is a well-known statistical concept, sufficiently simple to allow wide use and easy communication of results between researchers of diverse backgrounds.

On the other hand, from the mid-1980s, nonlinear approaches to analysis of brain signals, particularly EEG, are getting increased interest of researches who consider nonlinearity as an intrinsic property of brain dynamics. This nonlinearity is studied both in the univariate temporal structure [154] and the interdependence [131]. The fact that the nonlinearities may be only occasionally and weakly present in the biological datasets has previously motivated work on improvement and assessment of nonlinear interdependence detection methods [163, 172]. For a review of nonlinearity in EEG, see e.g. [168].

Importantly, the resting state fMRI signal temporal properties are different from those of EEG. Moreover, they have just recently become more intensively studied. Above all, apart from the neuronal dynamics nonlinearity, the haemodynamic nonlinearities are also known to affect the BOLD fMRI signal [40].

Notably, non-linearity of dependence between fMRI time series during resting state has been recently reported [104]. Use of non-linear measures of FC for the analysis of resting state data has been proposed [44, 121, 180], particularly including measures based on analysis of chaotic non-linear dynamical systems to analyse resting state data, suggesting that the assumption of linearity might be oversimplifying. The question arises, to what extent and in what context is it justified and beneficial to use non-linear measures of FC. This issue is also related to the question of justifiability of use of any indices motivated by the assumption of non-linear dynamical system as the underlying mechanism behind the resting state fMRI signal. These include temporal fractal dimensions or Lyapunov exponents of the BOLD signal. While the presented or similar investigation might inform us on this question as well, we focus on the topical question of suitability of linear correlation as a FC measure.

When linear correlation is used as a measure of FC, there are some implicit assumptions made. The first is that the information in the temporal order of the

samples can be ignored (both within each time series and the mutual interaction). While the extent of justifiability of this assumption deserves exploration of its own, we keep this interim assumption for the purposes of this chapter, not least in order to keep the comparison of linear correlation to nonlinear measures fair.

Accepting for now the assumption of no role of temporal order of samples, we ask if the instantaneous (zero-lag) dependence between the time series, expressed in the probability distribution $p(X, Y)$, is fully captured by the linear correlation $r(X, Y)$. We answer that this is true under the second important assumption: the assumption of bivariate Gaussianity of the distribution. Indeed, a multivariate normal (Gaussian) distribution is uniquely defined by its correlation - up to linear shifts and rescaling. To be more precise, a bivariate normal distribution is fully characterised by its mean $\mu = (\mu_x, \mu_y)$ and its 2×2 covariance matrix $\text{Cov}(X, Y)$ - if we allow for linear shifting and scaling, the remaining invariant parameter characterizing fully the distribution is indeed the correlation $r(X, Y)$. For a bivariate Gaussian distribution, the correlation also uniquely defines the mutual information shared between the two variables X, Y which can be computed as $I(X, Y) = I_{\text{Gauss}}(r) \equiv -\frac{1}{2} \log(1 - r^2)$.

On the other hand, when the Gaussianity assumption does not hold, the distribution cannot be fully described by the mean and covariance. More, possibly infinitely many, higher order moments need to be specified to determine the distribution. As the correlation is not sufficient to describe the dependence structure, the equation for $I(X, Y)$ above cannot hold in general. Interestingly, we can use the prominent properties of normal distribution to derive a useful lower bound on mutual information valid for a broad class of probability distributions. In particular, for a bivariate distribution $p(X, Y)$ with standard normal marginals $p(X), p(Y)$, it holds that $I(X, Y) \geq I_{\text{Gauss}}(r) = -\frac{1}{2} \log(1 - r^2)$, where the equality holds exactly for bivariate Gaussian distributions. This allows us to quantify the deviation from Gaussianity as the difference between the total mutual information of the two variables $I(X, Y)$ and the mutual information $I_{\text{Gauss}}(r) = -\frac{1}{2} \log(1 - r^2)$ that correspond to bivariate Gaussian distribution with the observed correlation r .

While there are many potential nonlinear FC measure candidates, mutual information holds a specific position among these for its generality. In theory, it is general enough to capture an arbitrary form of dependence relation between the variables without any prior model restrictions on its form. The properties of mutual information allow us not only to test the suitability of linear correlation

through probing the Gaussianity of the fMRI time series, but also to construct a quantitative estimate of connectivity information neglected by the use of linear correlation. This gives the amount of additional information available and bounds the potential contribution of non-linear alternatives over the Pearson correlation coefficient.

We implement the outlined ideas by comparing the total mutual information between the signals with the mutual information between the signals in surrogate datasets. These surrogates are generated in a way that preserves the linear correlation, but cancels any nonlinear information by enforcing a bivariate Gaussian distribution on the surrogate signal-pair. This approach allows us to both test and quantify the deviation from Gaussianity, providing a principled guide in judging the suitability of linear correlation as a measure of FC. The focus on bivariate Gaussianity as the crucial condition of suitability of use of linear correlation as FC index, along with the illustrative quantitative estimation of the deviation from Gaussianity by means of the mutual information neglected by linear correlation, are the two main contributions of this study to the discussion of fMRI FC methods. We apply the presented method to ROI-average time series obtained from resting state fMRI BOLD signal of healthy subjects, testing and quantifying the deviation from bivariate Gaussianity.

5.2 Material and Methods

5.2.1 Data

We use the preprocessed fMRI data from the study described in Chapter 4. In particular, for each of 15 healthy volunteers we had two 15 minute resting state fMRI scans. Each of these sessions produced 14 time series of 430 timepoints – one time series for each cortical ROI. The list of the ROIs and the method used to derive them was described in Chapter 4, including the preprocessing applied to the time series.

5.2.2 Analysis

The minimal information argument

As already mentioned in the Introduction, for a bivariate distribution $p(X, Y)$ with standard normal marginals $p(X)$, $p(Y)$, it holds that

$$I(X, Y) \geq I_{Gauss} = -\frac{1}{2} \log(1 - r^2), \quad (5.1)$$

where the equality holds exactly for bivariate Gaussian distributions. The inequality (5.1) stems from the fact, that normal distribution is the *maximum* entropy distribution for a given covariance matrix (or for a given correlation, as we assume without loss of generality that $\sigma(X) = \sigma(Y) = 1$). From the relation between mutual information and entropy ($I(X, Y) = H(X) + H(Y) - H(X, Y)$) it follows that mutual information of Gaussian distribution $I_{Gauss}(r)$ is then *minimal* from all distributions of given correlation r , under the assumption of fixed marginal entropies, which is true when the marginals have standard normal distribution. Note that the assumption of normality of the marginals is far less restrictive than it might seem. First, approximate data normality is commonly assumed in areas not restricted to fMRI FC analysis. More importantly, even if we find particular data deviated strongly from normality, any sample distribution can be monotonously transformed to match normal distribution before further statistical processing.

To assure precise non-Gaussianity estimates, we have indeed carried out this ‘normalization’ step. It consists in assigning the appropriate percentile to each value of a given variable and then replacing the original values of the variable by values corresponding to these percentiles in a standard normal distribution. Note that this normalization step does not affect mutual information between the time series.

For two discrete random variables X_1, X_2 with sets of values Ξ_1 and Ξ_2 , the mutual information is defined as

$$I(X_1, X_2) = \sum_{x_1 \in \Xi_1} \sum_{x_2 \in \Xi_2} p(x_1, x_2) \log \frac{p(x_1, x_2)}{p(x_1)p(x_2)},$$

where the probability distribution function is defined by $p(x_i) = Pr\{X_i = x_i\}$, $x_i \in \Xi_i$ and the joint probability distribution function $p(x_1, x_2)$ is defined analogously. When the discrete variables X_1, X_2 are obtained from continuous variables on a continuous probability space, then the mutual information

depends on a partition ξ chosen to discretize the space. Here a simple box-counting algorithm based on marginal equiquantization method [140] was used, i.e., a partition was generated adaptively in one dimension (for each variable) so that the marginal bins become equiprobable. This means that there is approximately the same number of data points in each marginal bin. We used a simple pragmatic choice of $Q = 8$ bins for each marginal variable [141].

For each session, we have computed the mutual information (MI) for each pair of regions, yielding a symmetric 14-by-14 matrix of MI values. To minimize bias of the MI estimates due to inevitable discretization and finite sample estimation, the MI values were further monotonously transformed to correct for these effects. This transformation map was generated using random samples from normal distributions with correlation ranging from 0 to 1 in 200 steps of 0.005. For each correlation value, 50000 such random bivariate samples with $N=430$ observations each were generated and the mean of their MI as computed by the equiquantization method was tabulated. As for bivariate Gaussian random distribution with correlation r the true MI is $I_{Gauss} = -\frac{1}{2} \log(1 - r^2)$, this tabulation allows approximate transformation of estimated MI to true bivariate MI.

Linear surrogate data

To compare the (total) mutual information to the portion of information conveyed in the linear correlation, for each dataset, 99 random realizations of multivariate time series preserving the linear structure but cancelling the nonlinear structure were constructed, and MI was computed for these surrogates. If the original time series dependence structure was Gaussian (and therefore fully captured by the linear correlation), the MI in the surrogates should not differ from the original MI, up to some random error. The alternative case should manifest itself as a decrease in the MI in the surrogates with respect to the original data.

The surrogates were constructed as multivariate Fourier transform (FT) surrogates [145, 138]: realizations of multivariate linear stochastic process which mimic individual spectra of the original time series as well as their cross-spectrum. The multivariate FT surrogates are obtained by computing the Fourier transform of the series, keeping unchanged the magnitudes of the Fourier coefficients (the amplitude spectrum), but adding the same random number to the phases of coefficients of the same frequency bin; the inverse FT into the time domain is then performed. The multivariate FT surrogates preserve the part of dependence which can be explained by a multivariate linear stochastic process.

The idea of comparing the MI of data to MI of ‘linear’ surrogates rather than directly to linear correlation of data has two aspects. First, it allows a direct quantitative comparison of the nonlinear and linear connectivity, while correlation and mutual information estimators have generally different properties. Second, generation of the surrogates allows direct statistical testing of the difference. However, this procedure generates 99 estimates of the linear MI for each parcel pair; one for each surrogate. While these are useful for hypothesis testing, for general presentation of the difference we use the mean value of these 99 values. In the following we refer to this as ‘Gaussian’ MI, and it actually closely estimates the MI of a bivariate Gaussian distribution $I_{Gauss}(r) = -\frac{1}{2} \log(1 - r^2)$, where r stands for the correlation of the two variables (see the close match in Figure 5.2, red and purple line). The ‘neglected’ MI is then estimated by the difference between data MI and the Gaussian MI: $I_{neglected}(X, Y) = I(X, Y) - I_{Gauss}(r)$.

5.2.3 Statistical tests

For each session and each region pair, non-Gaussianity was tested at $p = 0.05$ by comparing data MI against MI distribution of multivariate FT surrogates. To correct for multiple comparisons, the number of significant pairs in given session was then tested against the null hypothesis that the number of individual significant entries has a binomial distribution $B(n = 91, p = 0.05)$, where $n = 91 = \frac{14(13-1)}{2}$ is the number of all region pairs and $p = 0.05$ is the single entry false positive rate under condition of pure Gaussianity of the bivariate distributions.

As it may be argued that the assumption of pair independence is too lenient, but the exact level of dependence is difficult to establish, we also carried out group level tests. The percentages of significant pairs were compared by means of a paired t-test to the percentages of significant pairs obtained from *shadow datasets*. Each shadow dataset was created as a multivariate FT surrogate of normalized data of a given session, preserving only the linear structure of the dataset after normalisation of univariate marginals. Subsequently, each shadow dataset has undergone the same procedure as original data, including the initial normalization, generation of multivariate surrogates, computation of MI, and statistical testing of pair-wise MI against surrogates. In this way, we mimicked the full procedure using a ‘purely linear FC’ shadow dataset, accounting for any potential bias in the detection rate introduced by numerical properties of the algorithm. Apart from the percentages, we have also tested the mean neglected

information from data versus shadow datasets by mean of a paired t-test. The group-level tests used a $p = 0.05$ significance threshold, but we also report the attained significance level.

5.3 Results

5.3.1 Descriptive assessment

In descriptive terms, the data MI has proved very similar to the Gaussian MI (see Figure 5.1). In particular, averaging across all parcel pairs, the data MI ranged between 0.1 and 0.6 bits for different sessions, while the neglected MI was more than an order of magnitude smaller (-0.01 to 0.04 bits). Nevertheless, the estimates of neglected MI in data were generally positive (small negative values in only 3 out of 30 sessions), suggesting presence of some small but real nongaussianity. This was not the case for shadow datasets (ranging from -0.01 to 0.01 bits and generally symmetrically spread around zero).

Independently of the strength of coupling, the data MI was moreover typically within the range of surrogate MI, as illustrated by Figure 5.2. Here, each blue dot corresponds to MI of one parcel pair; the surrogate distribution is represented by red (light blue, green) lines for the mean (1st percentile, 99th percentile) of the surrogate distribution. Although the session with the most non-Gaussianity is depicted here, the distribution of computed MI for data and the corresponding shadow dataset (Figure 5.3) are almost indiscernible, with the former only slightly elevated. Also, apart from the random error due to MI estimation from short time series, which is shared by data and shadow data, both scatters follow well the theoretical prediction of dependence of MI on linear correlation ($I_{Gauss} = -\frac{1}{2} \log(1 - r^2)$, valid exactly under Gaussianity, purple line), which is closely approximated by the surrogate mean.

5.3.2 Statistical tests

The percentage of region pairs with significant non-Gaussianity was elevated in all but 3 sessions above the 5% expected under the null hypothesis (ranging from 0 to 35% of significant pairs in different sessions). If all the region pairs were considered independent this would constitute a significant percentage for 19 out of the 30 sessions considered (comparing to binomial distribution $B(n = 91, p = 0.05)$, where $n = 91 = \frac{14(14-1)}{2}$ is the number of all region pairs). As

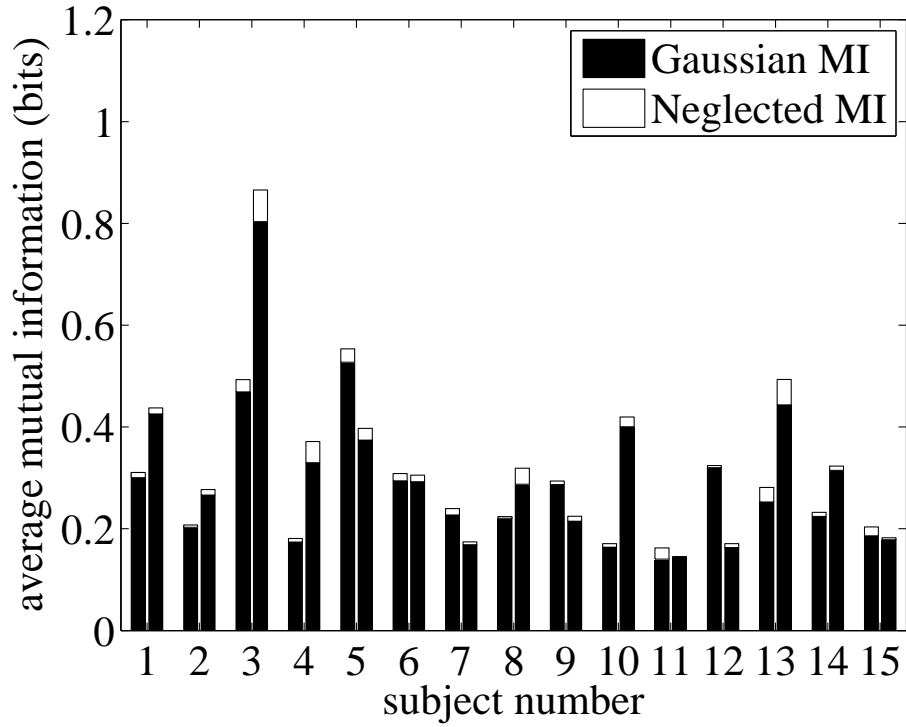


Figure 5.1: Comparison of the average Gaussian and neglected information. Each stack bar represents values for one session, averaged across all region pairs.

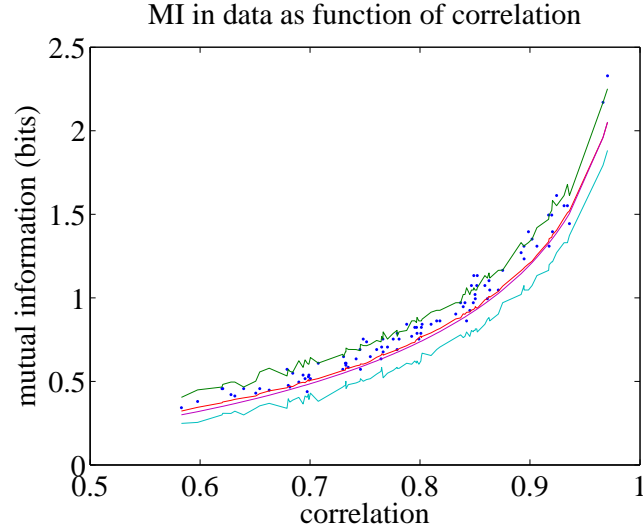


Figure 5.2: Mutual information as a function of correlation in an example dataset. The session with the most non-Gaussianity is depicted. Each blue dot corresponds to MI of one region pair; red (light blue, green) lines correspond to mean (1st percentile, 99th percentile) of the surrogate distribution. The purple line shows the theoretical mutual information of an exactly Gaussian distribution with the given correlation $I_{Gauss}(r)$ - it is not well visible as it closely matches the mean of the surrogate distribution.

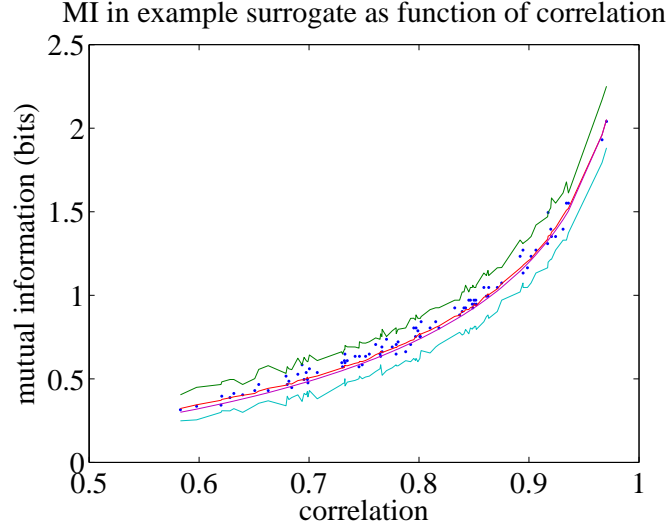


Figure 5.3: Mutual information as a function of correlation in an example surrogate. Each blue dot corresponds to MI of one region pair; red (light blue, green) lines correspond to mean (1st percentile, 99th percentile) of the surrogate distribution. The session with the most non-Gaussianity is depicted.

the assumption of independence of the pairs might be too lenient, yielding false positive results, we also carried out group level tests that confirmed the statistical deviation from Gaussianity. In particular, the counts of pairs with significant nonlinearity were significantly higher than similar counts obtained from shadow datasets, when compared on group level by means of a paired t-test ($t=3.91$, $df=29$, $p<0.005$).

Also, the neglected information in data averaged over parcel pairs was generally positive and on average had value 0.011 bits. On the other hand, the neglected information in the shadow datasets fluctuated around zero with mean of 0.002 bits. This difference was also clearly statistically significant ($t=4.09$, $df=29$, $p \sim 0.003$).

5.4 Discussion

The presented analysis revealed that the bivariate dependence structure of the fMRI BOLD regional time series is captured very well by linear correlation. Indeed, average mutual information was only several percent higher than the mutual information in surrogate data that contained only the linear part of the dependence. This gives an explicit and quantifiable argument for the intuitive choice of linear correlation as a measure of FC for fMRI time series.

Nevertheless, we have shown that there is a statistically significant contribution of non-Gaussian dependencies in the data, although the effect is so subtle that testing across many pairs or even across many sessions was needed to acquire sufficient power for such tests. The detection of non-Gaussian coupling is not surprising in the light of the fact that the dynamics of brain activity as well as the haemodynamic response of the vasculature contain many nonlinearities.

It is important to keep in mind that the observed deviations from Gaussianity might not only reflect a stationary non-Gaussianity in neuronal connectivity. In the presented framework, deviation from the null hypothesis could also be caused by nonstationarity of the signal. Vice versa, non-Gaussianity might lead to false detection of nonstationarity if the test assumes a linear Gaussian generating process as in [28]. Technically, a precise isolation of these two effects is extremely challenging. A practical consideration allows reasonably reconciling the two alternative interpretations: heavy nonstationarity could indeed increase the estimated non-Gaussianity, but it would also on its own invalidate the use of simple linear correlation, leading to the same conclusion about its suitability.

On another note, we also should not forget that we are working on the level of fMRI BOLD signal rather than with neuronal activations - both the Gaussian and non-Gaussian contributions to FC are likely to be affected to some extent by non-neural sources of signal variation. Nevertheless, in the end the estimation of whether this would play in favour of or against the use of linear correlation (and how much) seems to be entirely speculative at this point.

It can be argued that the amount of non-linear mutual information detected is likely to depend on the preprocessing of the data and the acquisition parameters. In this study, we have used standard acquisition and preprocessing. It may be argued that use of temporal filtering might have a Gaussianising effect not only on the univariate but also the bivariate distributions. To this end, we have recently analysed a different dataset without the use of low-pass filtering with similar results, presented in [81]. Also, temporal domain filtering is a generally accepted fMRI preprocessing step in order to remove artifacts. It could be argued that FC supported by the stop-band frequencies correspond mainly to non-neural artifacts [36].

Here we have applied the nongaussianity assessment on ROI-average time series. Importantly, under partial incoherence of voxel-level neural signals within the regions, some of the nonlinearity might be lost due to ROI-averaging. Therefore working on the level of temporal signal of spatial ICA components or single voxel

signal might in theory be more sensitive to nonlinearity. Although such FC is less easily interpretable and in the case of single voxels more prone to noise, exploring the dependence of the non-Gaussian contribution to FC on the time series extraction method is a subject of future work.

Also, the analysis presented in this chapter did not consider any time-lags. This is consistent with comparing to linear correlation as a measure of FC. The time lags would naturally have to be included in case that we were interested in assessment of causal or ‘effective connectivity’ measures (linear versus nonlinear). This defines another natural follow-up of the recent study.

Apart from some minor technical aspects such as normalisation of the time series, our approach differs from the previous probes into the potential of nonlinear fMRI FC such as [44, 104, 121, 180] mainly in that we explicitly focus on the bivariate Gaussianity rather than the linearity assumption as the condition of suitability of use of linear correlation as FC index. The deviation from this condition allowed us to quantify the potential available for arbitrary ‘nonlinear’ connectivity measures. This general interpretation is allowed by the use of a very general dependence measure – mutual information. In theory, this is able to capture virtually any form of statistical dependence. Of course, some practical limitations stem from the inevitably finite sample size, forcing us to summarize the results across ROI pairs and subjects. Last but not least, we provide an illustrative quantitative estimation of the deviation from Gaussianity by means of the mutual information neglected by linear correlation, that should give a theoretical upper bound on any improvement to be made by an arbitrary nonlinear connectivity measure. We note that an earlier attempt towards quantification of the FC ‘nonlinearity’ was made in [104], who reported also the explained variances by the higher order (nonlinear up to order 5) terms in the predictor set within a linear regression framework – although these variances were not corrected for the effect of extra number of regressors in the model, and also the considered model could not capture a more general form of dependence.

Despite the differences in the aims and methodology, our observations agree qualitatively with partial results of previous studies such as [104, 180] in that we also reject the specific hypothesis of stationary multivariate linear Gaussian process as the structure of resting state fMRI signal.

For clarity, we stress again that the above mentioned Gaussianity condition, rather than mere linearity of the process, is the key assumption for suitability of use of linear correlation. To fully acknowledge this consider that, as stated e.g.

in [12], a linear stationary process $(X_t)_{t \in \mathbb{Z}}$ is usually defined by

$$X_t = \sum_{j=0}^{\infty} \phi_j \epsilon_{t-j}; (t \in \mathbb{Z}), \quad (5.2)$$

where ϵ_t is i.i.d with $E[\epsilon_t] = 0$, $E[\epsilon_t]^2 < \infty$ and $\sum_{j=0}^{\infty} \phi_j^2 < \infty$.

While the definitions of linear process may differ in details, in most cases they are general enough to include processes with *non-Gaussian* distribution, which are then not fully described by their correlation structure. This may lead to some confusion, pointed out in [134], where it is shown that some widely used surrogate-based linearity tests such as those used in [180] are actually sensitive to non-Gaussianity. Careful examination of the assumptions of these tests reveals that the null hypothesis there is that the data come from a process that is equivalent to a linearly filtered white *Gaussian* noise. Of course, if the general definition of linearity is complemented by the condition that the generating noise is Gaussian, such a linear process is indeed also Gaussian, with probability distribution fully characterised by the mean and covariance structure. Discussion of the potential causes of rejection of the null hypothesis of ‘linearity’ in the FT-surrogate framework by processes that are still linear in the general sense of having the form (5.2) is also provided in [139].

For completeness, we note that linearity is also often discussed as an alternative to nonlinear, potentially chaotic deterministic dynamical systems. In this context caution is warranted with the interpretation of many ‘chaotic’ characteristics such as fractional correlation dimension or Lyapunov exponents when the underlying system might be of stochastic (non)linear nature rather than deterministic (non)linear dynamical system, and particularly when short time series such as those acquired from fMRI are being analyzed.

5.5 Summary

Quantification of the inter-regional synchronisation of spontaneous brain activity by functional connectivity measures is one of the most important tools for analysis of resting state fMRI data. The predominantly used functional connectivity index is the linear correlation of the local activity time series. Nevertheless, the use of nonlinear functional connectivity measures has been repeatedly advocated. In this chapter we investigated the question of suitability of the linear correlation for functional connectivity measurement. Our strategy was based around

the fact that linear correlation is an precise dependence measure when the studied variables follow bivariately Gaussian distribution, but when this condition is not satisfied, the non-linear dependence contribution tends to be neglected.

Thus, we assessed the suitability of linear correlation as a functional connectivity measure for fMRI time series by testing and quantifying the deviation from bivariate Gaussianity. To do this we used a fully general dependence measure — mutual information. The quantitative assessment revealed that the portion of mutual information neglected by using linear correlation instead of considering an arbitrary non-linear form of instantaneous dependence is relatively minor. Nevertheless, formal group-level test confirmed deviation of the data from the null hypothesis of the data being sampled from stationary linear Gaussian process. Overall we conclude that linear correlation of normalized data well captures the full functional connectivity: although existence of non-Gaussian contribution to bivariate dependences can be established, practical relevance of nonlinear methods trying to improve over linear correlation might be limited by the fact that the data are indeed almost Gaussian. We believe this provides a robust support for the use of linear correlation as fMRI functional connectivity measure.

Electrophysiological underpinnings of fMRI functional connectivity

In this chapter we present a specific perspective to studying the electrophysiological correlates of resting state fMRI signal. As discussed in section 3.5, previous studies usually correlated BOLD amplitude with instantaneous band-limited EEG power, typically in a chosen band of interest. Instead of working with the fMRI signal per se, we search for EEG correlates of *functional connectivity* as a major resting state feature. Recently, strength of FC within the DMN has been suggested to relate to the level of consciousness of a subject. Still, one of the current challenges lies in the biological significance of FC. Therefore, we carry out an inter-subject study of the electrophysiological correlates of the DMN FC, uncovering a network specific and functionally meaningful electrophysiological signature for decreased DMN FC. The results of this investigation have been published in the NeuroImage journal [78].

6.1 Material and methods

The data and preprocessing steps used in this study are partially identical with the data and preprocessing already described in the previous Chapter 4. For convenience, we detail here the full description of the material and methods including those that are shared with the previous chapter as well as the new data, preprocessing and analysis carried out solely for the purpose of this study.

6.1.1 Data acquisition

20 healthy volunteers (18 male, 2 female, age 18-35) were included after giving written informed consent. The study was approved by the Nottingham University Medical School ethics committee. Subjects were instructed to lie still with eyes closed and were awake and alert. This was confirmed behaviourally before and after the scan by an attending anaesthetist and by inspection of concurrent continuous EEG. A trained anaesthetist was present during the awake alert state as this served as a control condition in context of an unrelated study on the effects of midazolam sedation, not reported here. No EEG sleep features (K-complexes or sleep spindles) were detected.

Simultaneous EEG/fMRI were recorded on a Philips Achieva 3.0T MR scanner using an 8 channel head coil and standard multi-slice EPI ($TR/TE = 2100/35\text{ms}$, 64×64 matrix, $3.25 \times 3.25\text{mm}$ in-plane resolution, slice thickness of 3mm). A total of 35 contiguous transverse slices and 430 volumes were acquired. EEG data were collected inside the MR scanner from 30 scalp sites using sintered Ag/AgCl ring electrodes with built-in $5k\Omega$ resistors mounted on an electrode cap according to the international 10-20 system. Electrode FCz was used as a reference site. Two additional electrodes were placed below the left eye and on the upper back to monitor eyeblink and electrocardiogram respectively. Electrode impedance was maintained below $10k\Omega$. A BrainAmps MR high-input impedance amplifier (BrainProducts, Munich, Germany) specifically designed for use in high magnetic fields was utilised, placed in the scanner bore beside the head coil. Amplified EEG signals were transmitted via fibre optic cable to a recording computer placed outside the scanner room. The data were recorded with a pass-band of 0.016-250Hz at a sampling rate of 5kHz.

6.1.2 Data analysis

MRI

Standard pre-processing steps were applied, using FSL4.0 software package: non-brain voxel extraction, motion correction and high-pass 0.01 Hz frequency filtering. For each session, registration matrices to MNI template were computed and their inverses used to register MNI based ROIs to individual acquisitions. The ROIs were chosen to represent the main nodes of the DMN located in the “PC” (Precuneus/posterior cingulate), “mPFC” (medial prefrontal cortex), “ITPC” (left temporoparietal cortex) and “rTPC” (right temporoparietal cortex). To be

able to control for network specificity of the EEG/FC relation, four additional ROIs located at the main nodes of the Dorsal Attention Network (DAN) [49] were chosen.

The following procedure for deriving the ROIs was used: firstly, for each network a binary network template consisting of spherical regions (15 mm diameter) centred at significant network peaks was created. The peak locations were adapted from SI Table 3 of [118]. Talairach coordinates were converted to Montreal Neurologic Institute coordinates by using Matthew Brett’s `tal2mni.m` script, implemented in Matlab (<http://imaging.mrcbu.cam.ac.uk/imaging/MniTalairach>) [20]. For three of the ROIs, several mm adjustment was made, symmetrising the bilateral ROIs - the peak closer to the brain surface was replaced by symmetric copy of the contralateral peak, thus allowing the whole sphere ROI to be within brain volume.

Secondly, a template-matching procedure proposed in [73] was run on ICA decompositions of MNI-registered fMRI data to detect individual maps for each network. The ICA components were obtained using ICA as implemented in MELODIC (FSL4.0) [10], utilising built-in automated estimation of number of components. As a last step, group-based spatial mask for each network was generated by averaging the individual maps for the given network across individuals and binarizing the resulting average map by thresholding at $Z > 2$. The binary group-network masks were then split into several ROIs for subsequent long-range FC analysis. In particular, 4 ROIs were created for DMN: “PC” (Precuneus/posterior cingulate), “lTPC” (left temporoparietal), “rTPC” (right temporoparietal) and “mPFC” (medial prefrontal cortex) and 4 ROIs for DAN network (bilaterally dorsolateral prefrontal ROI and temporoparietal ROI). These 8 ROIs were complemented by 6 ROIs from three lower-level function networks, namely Auditory, Motor and Visual Network. The ROI for the lower-level functions were derived with analogous procedure with the difference that the centres of the template spheres were located at hemispherical peaks of anatomical maps in probabilistic Harvard-Oxford Atlas included in FSL 4.0: initial template for the Auditory Network was based on Heschl’s gyrus map, for the Motor Network on precentral gyrus map and for the Visual Network on intracalcarine sulcus map. For each of these three networks, only two ROIs were defined (splitting the final network mask into left and right half). All ROIs were checked for spatial overlap between them and the overlapping areas (although not constituting substantial part of any of the ROIs) were excluded from the

ROIs. The centre locations of the sphere templates are summarised in Table 4.2, the actual derived ROIs are shown in Figure 4.1.

Average time courses for each ROI of each subject were extracted (mean of all included voxels) and orthogonalised with respect to motion parameters and average white matter and CSF time courses (obtained from thresholded white matter and ventricular ROIs from the Harvard-Oxford Atlas included in FSL 4.0) to correct for nonspecific global signal fluctuations. A low-pass 0.1 Hz second-order Butterworth filter was applied to constrain the analysis to frequencies of interest and thus decrease the potential contribution of high-frequency content characteristic of some physiological confounds to the measured FC. FC was calculated by Pearson’s correlation coefficient for each pair of ROIs.

EEG

Brain Vision Analyzer software (Brainproducts) was used for correction of MR gradient and ballistocardiographic (BCG) artifacts as described elsewhere [3, 2]. Gradient artifacts were removed as implemented in Vision Analyzer software (BrainProducts, Germany) by subtracting an artifact template from the data, using a baseline-corrected sliding average of 20 consecutive volumes. Following scanner artifact removal, pulse artifact subtraction was applied. This procedure works analogously by averaging EEG signal synchronized to the ECG. The MR-denoised EEG data were referenced to common average. Subsequently, the EEG data was downsampled to 256 Hz and imported to EEGLab [43] (www.sccn.ucsd.edu/eeGLab), an open source toolbox running under the MATLAB environment. Segments contaminated by artifacts due to gross movements were removed following visual inspection (maximum total 3-4 minutes rejected from the 15 minute data in a given subject) and rare occasions of “bad” channels excluded from the analysis. ICA was performed on the continuous data using the infomax algorithm and components representing eye-blink, movement artifacts and residual ballistocardiographic (BCG) artifacts were removed from the data [167].

The power spectrum was calculated using a Fast Fourier transform on 4.2 second segments and averaged over the segments. The spectrum was also averaged over all EEG channels. Subsequently, the band-limited power in the 4 frequency bands, namely delta (1-4Hz), theta (4-8Hz), alpha (8-13Hz) and beta (13-30Hz) was computed by trapezoidal integration method. Note that while details of determining EEG band-limited powers generally vary between studies, in our

case the bands and the averaging over all channels were chosen in agreement with [118].

Prior to power computation, the residual scanner-related artifactual peak at 16.67 Hz was removed by a notch filter. Absolute band-powers may however not be best suited for between-subject comparison due to potential non-neural sources of between-subject variation of overall EEG power (e.g. skull thickness). Hence, we also computed relative band-powers (as a proportion of total EEG power of the given subject in the 1-30Hz band). At a second level of EEG analysis we have used the 95% spectral edge frequency (SEF 95) to assess the general slowing of EEG. SEF 95 index was computed for each subject according to [158, 162] from the cumulative power spectra as the frequency below which 95 percent of total EEG spectral power of the subject was located.

6.1.3 Statistical analysis

We subsequently tested the relation between EEG spectra and within DMN FC by means of a multiple linear regression with EEG band-power indices as explanatory variables and with the overall within DMN FC index as the dependent variable. The DMN FC index was computed for each subject as an average of the FC indices obtained for 6 possible ROI pairs within DMN - to increase measure robustness and decrease the multiple comparison problem. We report the explained cross-subject variance in the DMN FC index by the regression model and also partial correlations for each band. To test for network specificity of the EEG-FC link, the multiple regressions and partial correlations were also computed for the within DAN FC (average of 6 FC indices), and DAN-DMN FC (average across 16 FC indices representing FC between each of the 4 DMN sites with the 4 DAN sites). In addition, we tested explicitly the difference in the sets of EEG/FC regression coefficients obtained for DMN FC index and DAN FC (or DAN-DMN FC) index within a GLM framework, namely by means of the Chow test [30]. In brief, the Chow statistic is defined as:

$$\frac{(S - (S_1 + S_2))/k}{(S_1 + S_2)/(n_1 + n_2 - 2k)},$$

where S is the sum of squared residuals for a pooled regression model assuming common regression coefficients for both dependent variables, while S_1 and S_2 are sums of squared residuals for each of the regression model fitted separately for each dependent variable. Under the null hypothesis of regression coefficients

equality this statistic has Fisher distribution with k and $n1 + n2 - 2k$ degrees of freedom.

Further, we have used simple linear correlation to test the relation of FC indices to SEF 95 index. The difference of the correlation coefficients of SEF 95 with the two network FC indices was tested according to [17] by converting it to approximate t-statistic with $n - 3$ degrees of freedom by:

$$t = (r_{xy} - r_{zy}) \sqrt{\frac{(n - 3)(1 + r_{xz})}{2(1 - r_{xy}^2 - r_{xz}^2 - r_{zy}^2 + 2r_{xy}^2 r_{xz}^2 r_{zy}^2)}},$$

where y stands for the shared variable of the compared correlations (i.e. SEF 95), while x and z stand for the two other variables (the FC indices) and r_{ij} stands for correlation coefficient of variables i and j . For the relative band-powers only simple linear correlations are reported, since, due to inevitable multicollinearity, partial correlations were not well defined. All the results are reported as statistically significant based on a threshold $p < 0.05$; the p-values are also reported to give an idea of the significance level of the results. Conservative Bonferroni correction for multiple comparisons was used where appropriate; as indicated in the text.

6.1.4 EEG band-powers explorations

Before testing the link between EEG and FC, we explored the general distribution of the EEG power (see Figure 6.1) and the structure of relations between the EEG bands. The absolute EEG band-powers were all positively interrelated ($r = 0.18$ to 0.76). These correlations are similar to within subject temporal EEG band-power correlations reported e.g. in [39]. The relative powers showed positive correlation only between the alpha and beta band ($r = 0.47$), with negative relation for the other band pairs ($r = -0.74$ to -0.17).

6.2 Results

6.2.1 EEG predicts DMN functional connectivity

The linear regression model revealed significant link between the band-power EEG indices and the DMN FC index ($R\text{-square} = 0.70$, $p=0.001$). Two of the band-powers showed significant partial correlation to DMN FC index, namely delta ($r = -0.73$, $p=0.001$) and beta ($r=0.53$, $p=0.028$).

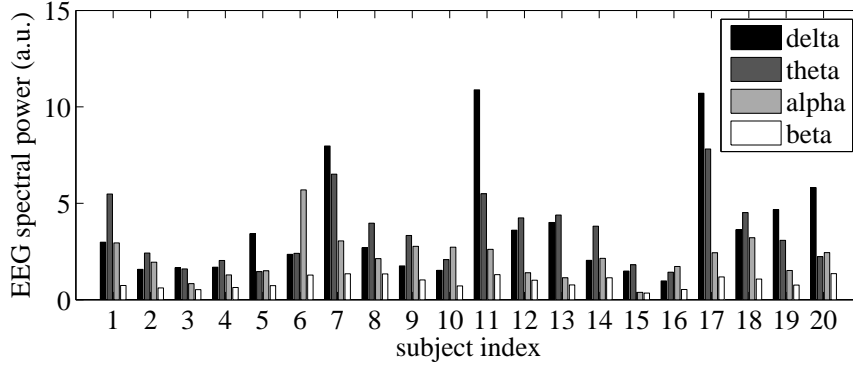


Figure 6.1: The general distribution of absolute EEG powers.

6.2.2 EEG and BOLD FC correlation is specific for within DMN FC

To assess whether the DMN-FC-related EEG profile was network specific, we repeated the analysis for within DAN FC and functional connectivity between DAN and DMN, with negative results. The multiple linear regression model was neither significant for the within DAN FC nor for the between network FC. Similarly, the partial correlations between the EEG bands and the DAN FC, DMN-DAN FC were all statistically insignificant. See Figure 6.2 for visual summary of the partial correlation results. Further, the explicit test rejected the equality of the set of regression coefficients for DMN FC and DAN FC ($p=0.023$), although did not reach significance for the comparison between DMN FC and the functional connectivity between DMN and DAN.

For relative band powers the results were quite similar. The multiple linear regression model showed significant link between relative EEG band-powers and DMN FC index ($R\text{-square} = 0.58$, $p=0.003$). The Pearson linear correlation analysis revealed significant correlation of three of the relative band-power EEG measures to overall DMN FC index. These were beta ($r = 0.69$, $p < 0.001$), alpha ($r = 0.55$, $p = 0.010$) and delta ($r = -0.61$, $p = 0.004$). These correlations remain clearly significant even after conservative Bonferroni correction for multiple comparisons. On the other hand, the relative theta power link to DMN FC proved insignificant ($r = -0.05$, $p = 0.83$). As noted in subsection 6.1.3, meaningful partial correlation could not be computed due to inherent multicollinearity of the relative EEG band-powers. None of the correlations of DMN FC and DAN FC with relative EEG band powers was found to be significant. See Figure 6.3 for summary of the correlation analysis for the relative band-powers.

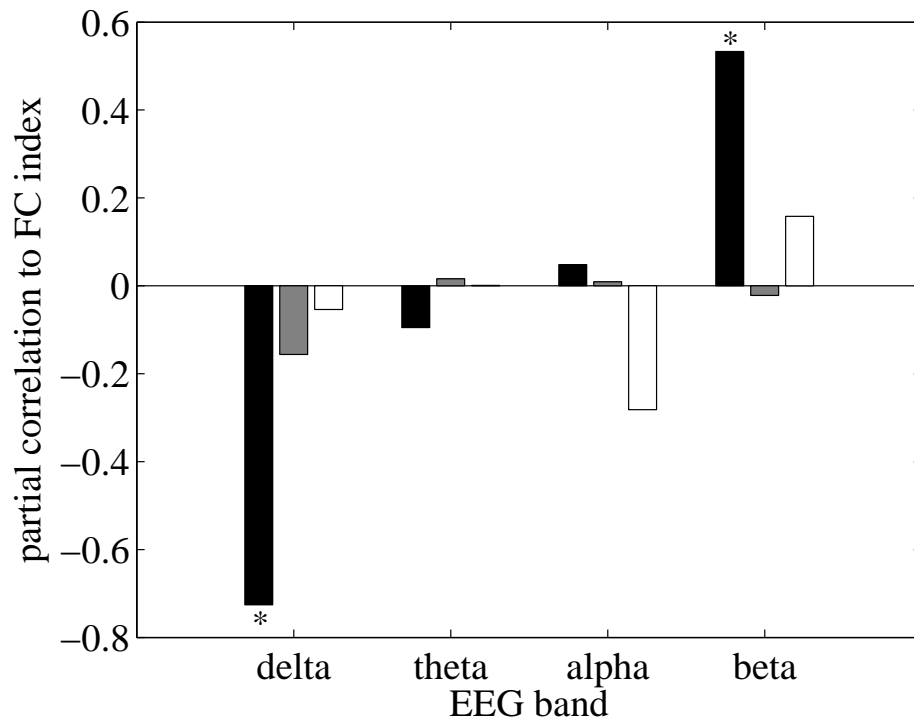


Figure 6.2: The relation between absolute EEG band powers and the FC in DMN, DAN and between them. Bars show partial correlation of each EEG band power to DMN FC index (black), DMN-DAN FC index (dark grey), DAN FC index (light grey). An asterisk (*) denotes statistically significant partial correlation ($p < 0.05$).

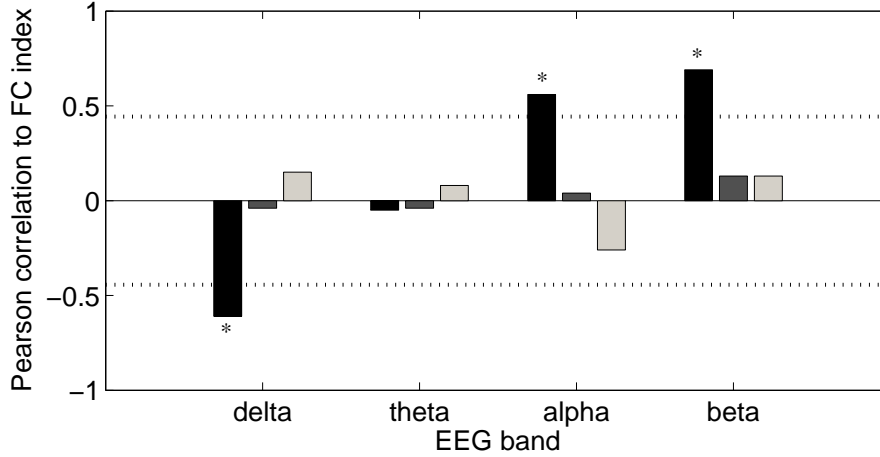


Figure 6.3: The relation between relative EEG band powers and the FC in DMN, DAN and between them. Bars show Pearson’s correlation of each relative EEG band power to DMN FC index (black), DMN-DAN FC index (dark grey), DAN FC index (light grey). Dotted line shows threshold for correlation significance ($p < 0.05$). An asterisk (*) denotes statistically significant correlations ($p < 0.05$). The difference between EEG/FC relation for DMN FC and both DAN FC and DMN-DAN FC was also found significant ($p < 0.05$) for delta, alpha and beta band.

6.2.3 DMN FC is correlated with EEG sedation index (spectral edge frequency)

As the overall pattern of correlation between DMN FC and EEG suggested that general slowing of EEG as seen in states of reduced alertness correlates with decreased DMN FC, we investigated a direct link between DMN FC and an established EEG-derived sedation index SEF 95. We found a strong direct correlation between SEF 95 and DMN FC ($r=0.70$, $p < 0.001$), whereas no correlations were seen between SEF 95 and FC within DAN or between DMN and DAN (Figure 6.4 left). Correlation between DMN FC and EEG differed significantly from that between DAN FC and EEG ($p = 0.002$) and DAN-DMN FC and EEG ($p = 0.02$).

6.2.4 Further exploration of EEG/FC network specificity

Summarizing the EEG into a single SEF 95 coefficient decreased the degrees of freedom for testing the network specificity of the detected DMN FC / EEG relation, increasing the statistical power sufficiently to show explicitly significant difference also for DAN-DMN FC. We also tested in exploratory fashion the SEF 95 - FC relation for the sensori-motor networks (auditory, motor and visual

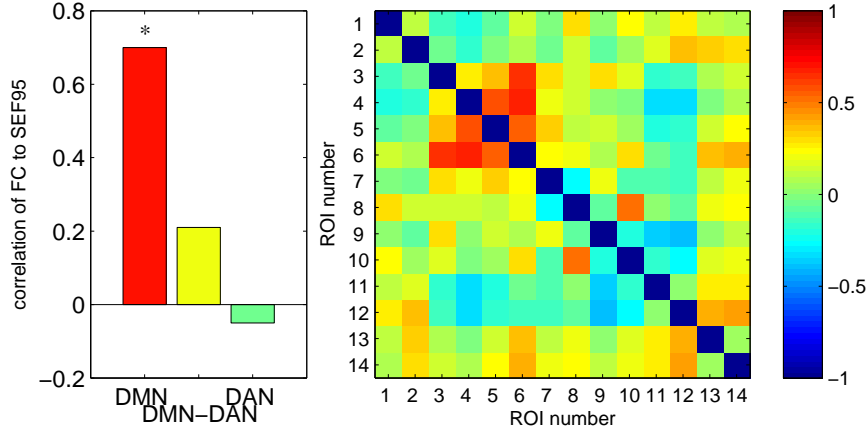


Figure 6.4: Correlation between fMRI FC and EEG spectral SEF 95 index. Left: correlation of SEF 95 to average FC within or across networks. An asterisk (*) denotes statistically significant correlation ($p < 0.05$). Right: full matrix of correlations of SEF 95 to between-ROI FC. The within - DMN connectivity shows distinct pattern of correlation to SEF 95 index. ROIs: 1: Auditory: left, 2: Auditory: right, 3: DMN: IPL left, 4: DMN: IPL right, 5: DMN: MPFC, 6: DMN: PCC, 7: DAN: DLPFC left, 8: DAN: DLPFC right, 9: DAN: PL left, 10: DAN: PL right, 11: Motor: left, 12: Motor: right, 13: Visual: left, 14: Visual: right.

network). We computed correlation of SEF 95 to FC within each of them as well as to FC averaged over all pairs between given sensori-motor network and DMN. All these correlations were found to be quite weak, ranging from -0.21 to 0.21. None of these correlations were significant, while all differed significantly from the strong positive DMN FC/SEF relation ($t = 2.60$ to 3.33 , $p = 0.004$ to 0.021).

A detailed picture of the relation between SEF 95 and pair-wise FC is shown in Figure 6.4 right. Note that since the FC was not averaged across pairs the results are generally more affected by noise than for the more robust network FC indices. Nevertheless, we can observe consistent positive relation between the summarising SEF 95 EEG index and DMN FC, compared to weaker and scattered effects within DAN, between DMN and DAN or elsewhere in the connectivity matrix.

6.3 Discussion

This study highlights a novel perspective of EEG/fMRI correlation by demonstrating interrelations between band-limited EEG power and inter-regional BOLD synchronisation. Using a between subject design, we found a strong link between

DMN FC and EEG band-powers with unique negative contribution of delta power and positive contribution of beta power. Furthermore, this association was found to be network-specific, as it was absent when studying FC between nodes of another cognitive network or FC between DMN and other sensori-motor or cognitive nodes. This underlines the notion of a specific EEG signature of the within DMN FC in addition to previously reported EEG signature of DMN activity.

Notably, the EEG band-powers explained 70% of the between subject variance of DMN FC using a multiple linear regression model. The strength of the association is intriguing in itself in light of the concern of spurious correlations in FC analysis due to scanner and cardio-respiratory induced noise. Our data provide evidence that careful data post-processing yields neurophysiologically meaningful data with most of the variance of DMN FC explained by electrophysiological findings that are robust against physiological factors known to confound FC fMRI analysis.

Subsequent analysis using relative instead of absolute EEG powers confirmed that the general EEG pattern associated with higher DMN FC consisted of increased high frequency and decreased low frequency EEG band power contribution. Univariate exploration in particular showed a close positive association between DMN FC and relative beta power and similarly strong negative associations with relative delta power. The ratio of explained variance goes from 48% if only beta is included slightly up to 58% when including the other frequency bands in a linear fit, reflecting that the variance explained by beta, alpha and delta is significantly overlapping. Taken together, the key feature of the detected EEG signature underlying increased within DMN connectivity can be summarised as a decrease in absolute and relative low-frequency (mainly delta) and increase in high-frequency (mainly beta) EEG power.

In other words, high delta and low beta EEG power was found to be associated with reduced individual DMN FC. Low-frequency dominant EEG is a common feature in states of reduced alertness such as drowsiness, sleep or sedation. This has been exploited to establish an EEG sedation index based on the spectral edge frequency such as SEF 95 (95th percentile of the EEG power spectrum) with higher frequencies reflecting increased alertness, lower frequencies increased sedation [158, 162]. Thus, the observed significant positive and network specific correlation between SEF 95 and DMN FC lends strong support that alertness positively covaries with DMN FC.

Several studies have been conducted to detect the robustness or sensitivity of DMN FC with respect to deviation from normal awake state. Notably, decrease of DMN FC has been reported in sedation [72] and, particularly a decreased mPFC - PCC FC, was observed in deep sleep [86, 155], although a recent light sleep study [106] did not detect significant DMN FC changes. Remarkably, a recent study [18] has shown that DMN FC is decreased in severely brain-damaged patients, in proportion to their degree of consciousness impairment, ranging from coma, to vegetative state, minimally conscious state and locked-in syndrome. DMN FC was therefore reported to be decreasing with decreasing level of alertness. Taken together our results are well in line with the principle of sedation underlying reduced FC within DMN. Our results extend previous findings by demonstrating that this general association can also be observed in healthy volunteers during relaxed wakefulness.

While both EEG slowing and decrease in DMN FC have been independently reported as markers of significantly reduced vigilance (such as due to sleep or sedation), in this study we have shown that these two observed phenomena are directly and strongly mutually related. Notably, a similar EEG/FC relation was recently reported across sleep-wake cycle within a single-subject 3 hour acquisition [86] (Fig. 3 therein).

It is important to bear in mind that comparison with previous combined BOLD and EEG studies is limited by the two distinctive methodological aspects of our approach. Firstly, we focused on EEG correlates of BOLD FC rather than absolute signal. Secondly, we explore between subject rather than intra-individual dynamic variance. To our knowledge, most of resting state combined EEG/fMRI studies reported within-subject correlation of EEG powers to fMRI signal time series, obtained either from voxels [39, 68, 107, 109, 128] or some weighted voxel combination such as an independent components [21, 118]. While some of these may use functional connectivity tools to obtain the representative fMRI signal (such as [118]), the single subject analysis presented in Figure 3 of [86] is actually a relatively isolated attempt to directly relate EEG measures to fMRI DMN FC.

Although the relation between network activity, fluctuation magnitude and connectivity is not yet fully understood [11, 50], there is growing evidence for congruence of changes in LFF amplitude and coherence elicited by alteration of the behavioural state [11, 99, 123]. For instance, [11] reported decrease in both amplitude and coherence of LFF in the visual cortex in eyes open fixation versus eyes closed resting state condition. Theoretically, the decrease in correlation could

be explained by a decreased signal-to-noise ratio due to lower signal amplitude given constant noise - even under constant inter-regional coupling.

To allow some approximate comparison of the observed FC EEG signature to EEG signatures of BOLD amplitude reported elsewhere, we will assume in the following that the changes in network activity, fluctuation magnitude and connectivity related to some EEG changes are in general congruent. Apart from the reports discussed above, some support for this stems from a theoretical speculation that while higher FC could be caused by higher BOLD temporal fluctuation under constant noise, this increased BOLD fluctuation correlated to particular EEG band power variability could in general relate to higher power in the particular EEG band. Under this model, the FC EEG signature would be inherited from the underlying BOLD EEG signature. Interestingly, conditions that suppress DMN activity were associated with reduced DMN connectivity if the experimental condition lasted long enough to allow FC estimation - see e.g. [54].

While it remains unclear to what extent BOLD connectivity is associated with BOLD amplitude it is of interest to compare our results to reported spatio-temporal correlations between EEG and BOLD amplitudes reports. Under the above assumption, the observed positive relation between high frequency EEG and DMN FC could be expected to translate into positive fMRI BOLD signal increases in DMN during transient increases of beta power (provided the effect of variation between subjects are similar to the effects of variation within subjects.)

Indeed, this would be in line with the finding that DMN activity was positively associated mainly with beta (and alpha) band [118]. Similarly, looking for dynamic correlates of 17-23 Hz activity in EEG Laufs et al. [108] reported fMRI activity in set of areas spatially resembling DMN. For low EEG frequencies, negative associations were generally noted with the DMN activity. Scheeringa et al. [156] reported a negative relationship between medial frontal BOLD signals and theta power rather than delta power. This partial discrepancy may constitute a true difference between the relationship of EEG with DMN activity and the relationship of EEG with connectivity within the DMN or might be due to the fact that the authors examined the relationship between BOLD signal and the theta power in only a single fronto-central component of EEG activity, derived by independent component analysis. Also note that while we reported analysis using partial correlations (or relative spectral powers), absolute theta power did indeed have negative correlation ($r=-0.48$) to DMN FC in our sample - likely

due to its correlation with absolute delta power.

While covariance of sedation/alertness is the most likely interpretation of the demonstrated EEG signature of DMN FC, we cannot exclude further contributions from inter-subject differences in mental state or predisposing personality. In this context it is worth mentioning that a positive correlation was noted between resting state connectivity within the salience network and inter-individual anxiety ratings [159]. Of note, an EEG profile similar to the one observed in our study has recently been found to correlate with both trait and state anxiety [100]. In addition, the link between EEG and state-anxiety has therein been shown to be mediated by the trait anxiety. To study such potential influences combined behavioural, electrophysiology and functional connectivity MRI studies will be required that will be further challenged by potential moderation or mediation effects between alertness on anxiety.

Functional disconnection is thought to be a key pathomechanism in a number of diseases and can be reflected in altered oscillatory brain activity. A detailed discussion of similarities between disease signatures of oscillatory brain activity as indexed by resting state EEG and resting state BOLD FC is beyond the scope of this article. Nevertheless, one disease that is characterised by a resting EEG signature similar to the one we found to predict low DMN FC is Alzheimer’s disease [103] that is also associated with reduced DMN FC [73]. The precise mechanism of altered EEG and DMN FC in Alzheimer’s disease remains unclear, but may be linked to cholinergic deficit leading to impaired neocortical arousal as a result from degeneration of the nucleus basalis (NB). In fact, lesioning of NB in rats increases delta EEG activity [25], and delta rhythms in Alzheimer’s disease patients were found correlated to impaired memory [7] and ‘normalised’ during cholinergic treatment [8].

6.4 Summary

Integrating the findings from different neuroimaging methods is one of the key tasks of current neuroscience. In particular, fMRI observations may due to their indirect nature benefit from comparison to electrophysiological measurements which may be argued to be in their nature closer to the underlying neuronal activity. Such comparisons of instantaneous BOLD fMRI *activity* with instantaneous band-limited power in crucial EEG frequency bands have been already reported in the literature.

In this chapter, we presented an approach opening a new perspective to understanding the EEG-fMRI relationship. In particular, we link the EEG signature represented by the powers in the main frequency bands to the strength of fMRI *functional connectivity*. Using this approach we demonstrate a significant and network specific relation between decreased functional connectivity within the DMN and a specific EEG pattern characterised generally by slowing of EEG. This provided a direct link between DMN FC and EEG signature that has been previously associated with altered states of alertness and anxiety trait and state levels providing a bridge between separate branches of research into neural correlates of brain state. This gives further evidence that properties of DMN FC may serve as index of mental state in relation to physiological and pathological variations in alertness/sedation. To the best of our knowledge this study provides the first direct evidence that DMN FC in awake state has a neurophysiological correlate thus validating its biological relevance, and providing a neurophysiological framework to interpret altered DMN FC.

Further data-driven investigations

In this chapter some questions are discussed that arise from the work described in Chapters 4 and 6 . More specifically, we ask to what extent does the mediatory effect of motion observed in LFF play role in also in the functional connectivity results.

7.1 Does motion mediate the link between EEG and DMN FC?

In Chapter 4 we discussed the mediation effect that head movement could have on the apparent increase of LFF power during sedation in comparison with baseline resting state condition. This effect was indeed shown significant for all five of the tested networks under standard preprocessing, and remained significant in at least two of these networks when additional global mean correction was applied. In addition, we have shown that motion amount correlates with the observed LFF power across subjects even within the normal resting state condition. On the other hand, in Chapter 6 we have discussed the inter-subject correlations between the EEG signature and fMRI functional connectivity within (or across) networks. Following the same logic as in our investigation of the motion effect on LFF in Chapter 4, one can naturally ask whether the correlation between functional connectivity and electrophysiological signatures as observed in Chapter 6 is genuine, or also mediated by some EEG-correlated head movement change. We set out to test this scenario in this section. As in Chapter 4, we follow the method of causal steps by Baron and Kenny [9].

7.1.1 Methods

Let us first recapitulate the general data situation. We are using the data described in detail in the Methods sections within the previous chapters. In particular, the 20 baseline sessions of resting state fMRI, the motion estimates obtained through the motion correction algorithm and the simultaneously acquired EEG are used in this analysis. Therefore, for each subject we have the head movement estimate M representing the mean relative displacement during the scan; four EEG measures - total powers in the delta, theta, alpha and beta band - represented by an explanatory (multi)variable X ; and the FC measures represented by the dependent variable Y - for initial analysis we will use the 3 average FC indices for connectivities within-DMN, within-DAN and between DMN and DAN. See Figure 4.2 for a graphical representation of the mediation model.

As discussed in Chapter 4, establishing mediation of the $X \rightarrow Y$ relation by the variable M requires assessment of four basic points:

- the existence of a link from X to M
- the existence of a link from M to Y
- the existence of a link from X to Y in a model that does not include M and finally
- documenting a significant decrease or indeed a complete disappearance of the link from X to Y after inclusion of M in the model.

To check the links between motion and EEG and FC respectively, we use simple linear correlations. As there are 4 EEG bands of interest, we also fit a multiple regression model with the EEG bandpowers as independent variables and motion as the dependent variable.

To assess the link between EEG and FC when motion is included, we use a multiple linear regression model, which allows a direct comparison to the results of Chapter 6, obtained without motion in the model. For all statistical tests, we use the uncorrected threshold $p = 0.05$ for statistical significance, unless stated otherwise. Due to the multivariate nature of the investigation, in this Chapter we, for simplicity, use a classical parametric approach for testing linear links rather than bootstrapping used in chapter 4. However, reviewing the evidence we can argue that the potential slight differences in the p-value estimates are not critical for the overall results.

	delta	theta	alpha	beta	full model
R	0.34	0.31	-0.24	-0.04	0.49
R-squared (unadjusted)	0.11	0.10	0.06	0.00	0.24
p-value	0.29	0.15	0.30	0.99	0.36

Table 7.1: Link between EEG and motion. The observed correlations were not significant.

	DMN	DMN-DAN	DAN
R	-0.28	-0.01	0.07
R-squared (unadjusted)	0.08	0.00	0.01
p-value	0.23	0.96	0.77

Table 7.2: Link between motion and functional connectivity in cognitive networks. The observed correlation was not significant.

7.1.2 Results

Link between motion and EEG None of the correlations between motion estimate and the 4 EEG absolute powers was significant, neither was the multiple linear regression fit. See the Table 7.1 for details of the correlations.

Link between motion and fMRI functional connectivity None of the correlations between motion estimate and the 3 summary FC indices was significant; see the Table 7.2 for details of the correlations. Figure 7.1 shows a more detailed picture of the relation between motion and functional connectivity across the baseline scans, suggesting a trend of overall positive relation between motion and functional connectivity within and between the lower level networks, but not so much for the networks of our primary interest in Chapter 6.

Link between EEG and fMRI functional connectivity corrected for motion The relation between EEG absolute powers and the FC indices did not significantly change with the inclusion of the motion estimate as an explanatory variable in the regression model for FC, as can be clearly seen from visual comparison of the partial correlation coefficients in Figure 7.2. The minuteness of the changes is confirmed by Figure 7.3, where the partial correlations are provided between each of the ROI-pair and EEG-band both without and with motion included in the regression model.

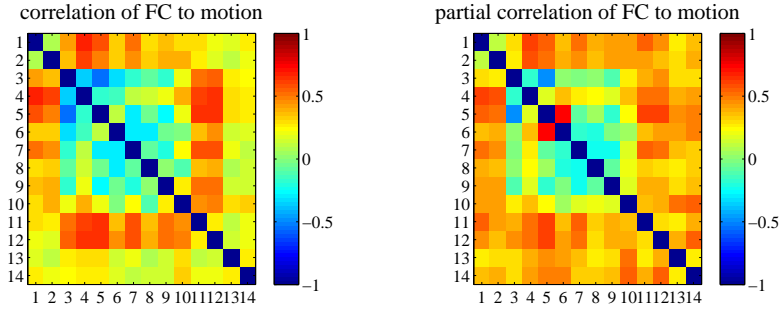
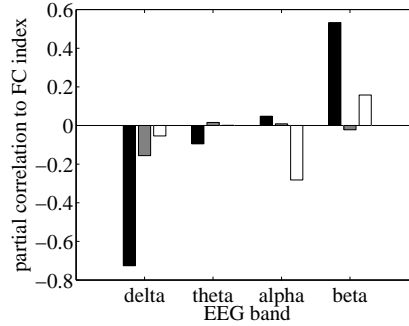
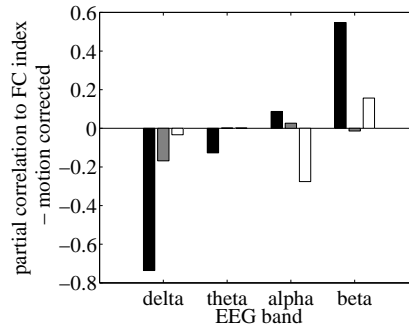


Figure 7.1: Correlation of motion and functional connectivity. Left: linear correlation of motion to FC between given pair of ROIs. Right: Partial correlation of motion to FC, correcting for 4 EEG absolute powers that were included in the multiple regression model. ROIs: 1: Auditory: left, 2: Auditory: right, 3: DMN: IPL left, 4: DMN: IPL right, 5: DMN: MPFC, 6: DMN: PCC, 7: DAN: DLPFC left, 8: DAN: DLPFC right, 9: DAN: PL left, 10: DAN: PL right, 11: Motor: left, 12: Motor: right, 13: Visual: left, 14: Visual: right



(a) Motion not included



(b) Motion included

Figure 7.2: The relation between the absolute EEG band powers and the FC within DMN, DAN and between them. Bars show partial correlation of each EEG band power to DMN FC index (black), DMN-DAN FC index (dark grey), DAN FC index (light grey). Left: motion not included in the inter-subject regression model. Right: motion included in the inter-subject regression model.

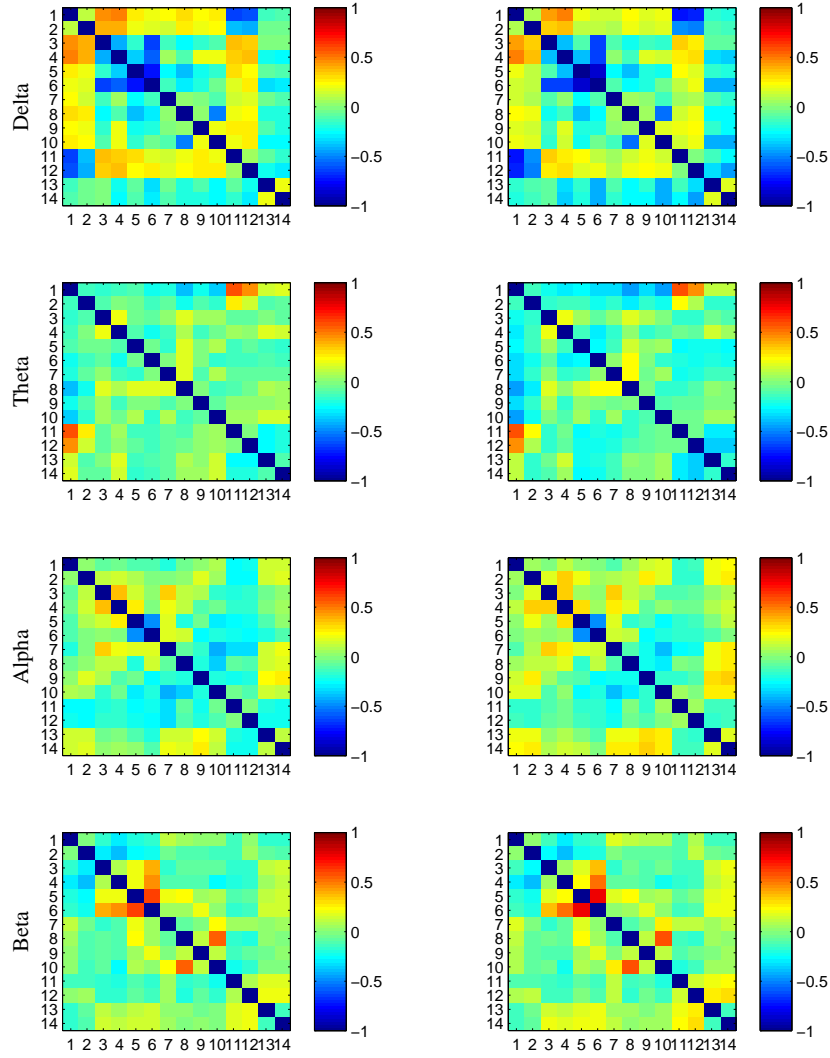


Figure 7.3: The relation between the absolute EEG band powers and FC. The colour-coding shows partial correlation of each absolute EEG band power and FC for given pair of ROIs. Left: motion not included in the model. Right: motion included in the inter-subject regression model. Numbering of the ROIs: 1: Auditory: left, 2: Auditory: right, 3: DMN: IPL left, 4: DMN: IPL right, 5: DMN: MPFC, 6: DMN: PCC, 7: DAN: DLPFC left, 8: DAN: DLPFC right, 9: DAN: PL left, 10: DAN: PL right, 11: Motor: left, 12: Motor: right, 13: Visual: left, 14: Visual: right

7.1.3 Discussion

Overall, we did not find a significant mediating effect of head motion on the observed EEG-FC relation. This can be mainly attributed to the fact that the head-movement was barely correlated to EEG powers. Also, within the networks of most interest for our analysis, i.e. the default mode network and the dorsal attention network, the observed functional connectivity was in general independent of head-movement.

Nevertheless, we have observed quite wide-spread correlation of head-movement to functional connectivity with, within and between the lower-level networks, namely auditory, motor and visual network. Assuming that this represents an artifactual correlation due to uncorrected artifacts, this would suggest that studies comparing functional connectivity across conditions that make subjects prone to different levels of head-motion may detect functional connectivity changes spuriously biased towards increases in functional connectivity. We investigate this hypothesis in more detail in the following section.

7.2 Could motion mediate sedation-induced changes in functional connectivity?

As shown in the previous section, in the baseline resting state there was a correlation between motion and functional connectivity, particularly affecting connections between networks/areas that are typically weakly connected. Assuming that is a spurious FC due to imaging artifacts, we hypothesise that this could lead to spurious increases in FC in sedation condition, particularly of those connections affected even during normal resting state. As this is already very data-driven investigation with many degrees of freedom, we refrain here from formal statistical testing, although we follow the general steps for testing mediation. These consist in establishing four points:

- existence of the total effect of sedation on FC
- existence of the effect of sedation on motion
- existence of the effect of motion on FC
- non-existence of the direct effect of sedation on FC when motion is included in the model (or significant weakening of the effect with respect to the total effect)

To establish the first point we show that there is indeed a sedation-induced increase in FC. This is illustrated in the upper subfigure of Figure 7.4.

Secondly, the effect of sedation on increased motion has been already shown in the Chapter 4. It therefore remains to inspect the effect of motion and the direct effect of sedation.

These are documented in the middle and bottom plot of Figure 7.4. The middle subfigure shows ab , i.e. the effect of motion b already multiplied by the scalar value a of effect of sedation on motion. Therefore, the middle subfigure shows the change in FC in sedation that was mediated by motion. The bottom subfigure then shows the direct effect of sedation on FC.

7.2.1 Discussion

The comparison of the middle and bottom subfigure of Figure 7.4 documents that most of the sedation-induced increase in FC, apparent mainly in between-networks connections, can be attributed to increased motion. In other words, when motion estimate is included in the inter-subject model as an explanatory variable for the FC, the sedation-related FC change is comparably weaker and less general. For motion-corrected FC, only one of the ROI-pairs showed motion-corrected FC change significant ($p = 0.014 < 0.05$, no multiple comparison correction), which is clearly not very reliable with as many as $(14 \times 13/2) = 91$ unique ROI-pairs in the matrix.

The fact that sedation-induced FC increases happen generally between networks, where the FC is very low during normal resting state, gives a strong argument for believing that these are artifactual. This is generally supported by the fact that these links correspond to the links that have FC correlated to motion even during baseline, see Figure 7.1. Nevertheless, as already discussed in Chapter 4, interpretation of results of mediation analysis needs to carefully take into account possible deficiencies of the model. In particular, we cannot fully disprove the alternative interpretation, that this specific pattern of increased connectivity in the between-network links is a true neuronal effect that for some reason well correlates with the amount of head motion. For instance, one could hypothesize that there exists some ‘global synchrony state’ in which the subjects also move more. Moreover, this ‘global brain synchrony’ state would be supported by sedation, but not exclusively as we have observed that the amount of motion remains a better marker of being in this state than the fact that the subject is

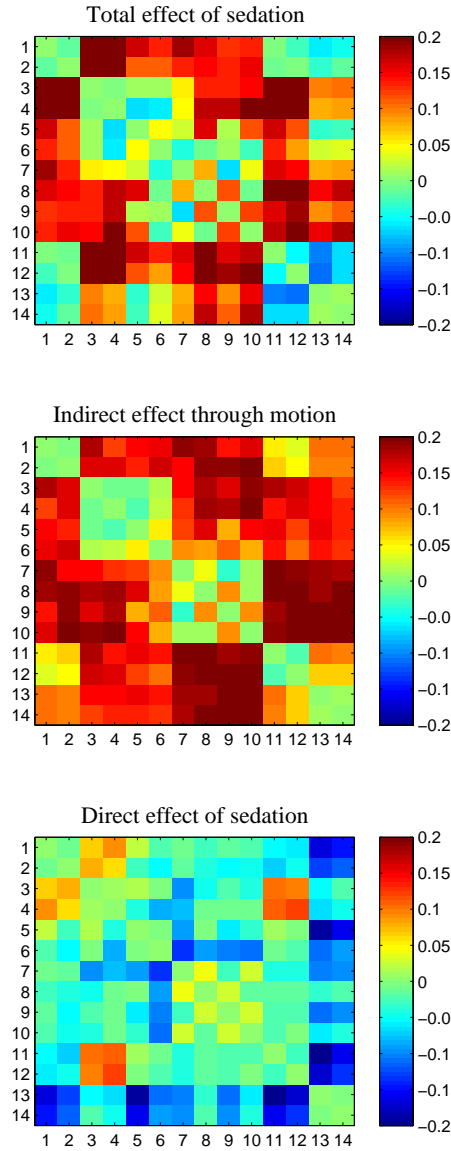


Figure 7.4: The sedation effect on FC. Top: total effect of sedation. Middle: indirect effect of sedation on FC, mediated by motion. Bottom: direct effect of sedation on FC, when controlling for motion increase. Numbering of the ROIs: 1: Auditory: left, 2: Auditory: right, 3: DMN: IPL left, 4: DMN: IPL right, 5: DMN: MPFC, 6: DMN: PCC, 7: DAN: DLPFC left, 8: DAN: DLPFC right, 9: DAN: PL left, 10: DAN: PL right, 11: Motor: left, 12: Motor: right, 13: Visual: left, 14: Visual: right

sedated, or than the EEG signature of the brain.

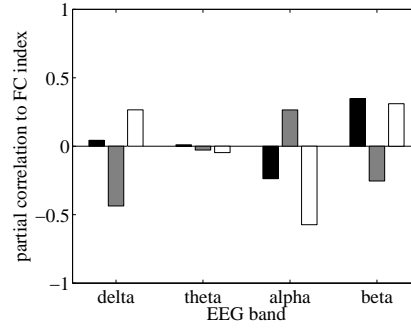
Another question is why we did not observe sedation-induced changes in DMN FC. It seems intuitive that sedation should induce changes in EEG, and these should be reflected in the DMN FC (according to our findings from baseline). While we did not investigate the sedation effects on EEG in detail, we can summarise as follows.

There indeed was a sedation induced effect on EEG, generally characterised by slowing of the EEG - there was a strong increase in delta and theta power and decrease of the SEF 95 index. However, we propose at least two reasons why this did not lead to a detectable decrease of DMN FC. Firstly, the EEG/FC link might be altered by the sedation. This is to some extent the case as can be seen by comparing Figure 7.2 and 7.5. While the EEG/FC relation is weaker and less specific than as described for the baseline resting state, we have still observed a significant ($r = 0.54, p < 0.05$, uncorrected) correlation of decreasing SEF 95 index characterising overall EEG slowing with decreased DMN connectivity (see Figure 6.4 in Chapter 6 for results in baseline). Similarly to the situation in baseline, this SEF 95 link was not significant for the DAN FC or the FC between DAM and DAN. Second explanation for not observing an EEG-mediated, sedation-induced DMN FC increase is that there might be *another* process with different or negligible EEG signature taking place, compensating for the extrapolated EEG-related FC decrease.

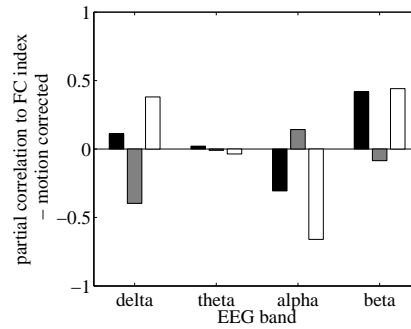
We have already mentioned in Section 3.4 that the reported changes of FC in reduced states of vigilance show mixed results with both increases and decreases being reported. Moreover, most of the studies focus on FC within networks. While we cannot relate our observations directly to a particular published study, they generally suggest that FC changes might be masked or escalated by motion changes, but this effect should not be prominent in within-network FC.

7.3 Summary

Functional connectivity is a key characteristic of spatiotemporal patterns of spontaneous brain activity. Therefore, the study of its relation to other variables describing the brain state such as electrophysiological or behavioural indices has an important place in neuroscience. The detected confounding role of non-neuronal processes, in particular head-motion, on the low-frequency fluctuation power estimates (see Chapter 4) motivated us to explore the confounding role of motion



(a) Motion not included



(b) Motion included

Figure 7.5: The relation between absolute EEG band powers and the FC in DMN, DAN and between them - sedated subjects. Bars show partial correlation of each EEG band power to DMN FC index (black), DMN-DAN FC index (dark grey), DAN FC index (light grey). Left: motion not included in the inter-subject regression model. Right: motion included in the inter-subject regression model.

in functional connectivity studies.

Firstly, we used the mediation model introduced in Chapter 4 to test the potential mediating effect of head-motion on the link between EEG and functional connectivity within the default mode network. This investigation did not detect any significant mediation, supporting the biological relevance of the observed EEG-FC pattern described in Chapter 6. Subsequently we asked if motion could mediate sedation-induced changes in functional connectivity. Indeed, we have observed sedation-related changes in functional connectivity, mainly between the studied networks, and these changes were attributable to the increase in involuntary head-motion in the sedation condition. Overall this chapter added further evidence and provided deeper discussion of the nature of the relation of functional connectivity changes to electrophysiological and behavioural markers of brain state.

Part III

Modelling Spontaneous Brain Activity

Models in computational neuroscience

On the quest for understanding brain function, researchers almost inevitably formulate models of neural behaviour. These models differ in many aspects, such as the level of formalisation, level of detail or biological realism. In this chapter we introduce some important representatives of models of single neurons and neuron populations, which are formalised through a full mathematical description. Models like those outlined are widely used in modern computational neuroscience. In the last section we describe recent application of such models in the study of spontaneous brain activity features observed in resting state fMRI studies, to the analysis of which the previous part of this Thesis was dedicated.

Following [92], before we move towards biologically more realistic models, sometimes called *biological neuron models*, we make an important theoretical as well as historical remark. In particular, we shall mention one example of a crude abstraction from the function of a real neuron, which is nevertheless the basic model used in so-called *artificial neural networks*. This simplified model is widely known as the McCulloch-Pitts model, as it was originally introduced in the classical paper “A Logical Calculus of Ideas Immanent in Nervous Activity” [124] by Warren McCulloch and Walter Pitts.

The model describes a network of neurons by a discrete-time dynamical system for the activity $z(i, t)$ of neuron i at time $t \in \mathbb{Z}$ by

$$z(i, t + 1) = S \left(\sum_j w_{ij} z(j, t) - \epsilon_i \right),$$

where w_{ij} is a weight of connection from neuron j to neuron i , S is a step function

and ϵ_i is a ‘threshold’ of neuron i . In other words, this equation only represents the fact that neural inputs are summed with given weights and converted to binary output based on whether the sum exceeds the threshold.

8.1 Hodgkin-Huxley model

Contrary to the model outlined above, the Hodgkin-Huxley model bears much more resemblance to biology. In fact, it is based on a biophysical description of the ion currents through the neuron cell membrane. Thus, it offers insight into the generation of the action potential, a stereotypical depolarisation of the membrane that is triggered by sufficient stimulation of the neuron by incoming currents. Note that the Hodgkin-Huxley model is still a very simplified neural model, and focuses only on one, although crucial, aspect of the real neural cell — the generation of action potential due to the action of voltage-gated ion channels. On the other side, it does not incorporate any detailed notion of receiving input from other neurons, its propagation through the dendritic tree modulated by spatial properties, or the propagation of the action potential along the axon and the mechanism of synaptic action through release of neurotransmitters to the synaptic cleft. We refer the reader to more specific textbooks on neural modelling such as [101] for more information on models of these processes.

In 1952, Hodgkin and Huxley performed a set of recordings from a squid giant axon, concluding that the action potential is mainly driven by sodium and potassium currents. Based on these recordings, they proposed what may appear to be a conceptually straightforward biophysical model built from the first principles. In particular, the crucial ordinary differential equation expresses the conservation of electric charge via the current balance equation

$$C \frac{dV}{dt} = -F + I_s + I, \quad (8.1)$$

where C stands for the membrane capacitance, V for the membrane potential (voltage difference across the membrane), F the membrane current, I_s the sum of the external synaptic currents and I is the current injected directly in experimental conditions. The membrane current F is further a function of the membrane potential V as well as gating variables m, n and h :

$$\begin{aligned}
 F(V, m, n, h) &= I_K + I_{Na} + I_L = \\
 &= g_K n^4 (V - V_K) + g_{Na} m^3 h (V - V_{Na}) + g_L (V - V_L).
 \end{aligned} \tag{8.2}$$

The membrane current is therefore a sum of the currents due to potassium and sodium ions, and all the other currents (including chloride ion) that are lumped into so-called leakage current I_L . The currents depend on the difference between the actual membrane potential and the respective reversal potentials (V_K, V_{Na}, V_L) for potassium, sodium and leakage, and further on their respective conductances. In the case of sodium and potassium currents, the conductance is further dependent on the state of gating variables m, n, h . These represent the fraction of channels in an open state. Their evolution in time, dependent on the membrane voltage, is described by:

$$\frac{dm}{dt} = \alpha_m(V)(1 - m) - \beta_m(V)m, \tag{8.3}$$

$$\frac{dn}{dt} = \alpha_n(V)(1 - n) - \beta_n(V)n, \tag{8.4}$$

$$\frac{dh}{dt} = \alpha_h(V)(1 - h) - \beta_h(V)h. \tag{8.5}$$

The model is finalised by the parameter values $C = 1\mu F cm^{-2}$, $g_L = 0.3 mS cm^{-2}$, $g_K = 36 mS cm^{-2}$, $g_{Na} = 120 mS cm^{-2}$, $V_L = -54.402 mV$, $V_K = -77 mV$, $V_{Na} = 50 mV$ and by the following equations provided by fitting the model to experimental data:

$$\alpha_m(V) = \frac{0.1(V + 40)}{1 - e^{-0.1(V + 40)}} \quad \beta_m(V) = 4.0e^{-0.0556(V + 65)} \tag{8.6}$$

$$\alpha_n(V) = \frac{0.1(V + 55)}{1 - e^{-0.1(V + 55)}} \quad \beta_n(V) = 0.125e^{-0.0125(V + 65)} \tag{8.7}$$

$$\alpha_h(V) = 0.07e^{-0.05(V + 65)} \quad \beta_h(V) = \frac{1}{1 + e^{-0.1(V + 35)}}. \tag{8.8}$$

To investigate the behaviour of the model, we carry out simulations and bifurcation analysis using the XPPAUT software [46]. The same software was also used for bifurcation analysis in the other sections of this chapter.

Under no external current input ($I = 0$), the Hodgkin-Huxley neuron relaxes to a resting potential $V \sim 75 mV$ - in other words, it is attracted to a stable fixed point of the dynamics of the model. As the current input increases, the

value of the voltage of the fixed point increases until at $I \sim 9.78$ this stable fixed point undergoes a Hopf bifurcation, coalescing with an unstable limit cycle to give rise to an unstable fixed point (see bifurcation diagram at Figure 8.2). The bifurcation structure between $I = 6$ and $I = 10$ is actually more complicated, including a small window of chaos [75]. Nevertheless, under constant current input $I \gtrsim 9.78$, an unstable fixed point and a stable limit cycle are the only invariant sets. Therefore the Hodgkin-Huxley model neuron tends to produce a periodic chain of action potentials. Such a ‘spike train’ for $I = 10$ is shown in Figure 8.1. Note that the Hodgkin-Huxley model starts firing with a non-zero frequency. This type of behaviour is sometimes called *Type II* neural excitability. A plot of the dependence of the firing frequency f on the current input (or *drive*) I is shown in Figure 8.3.

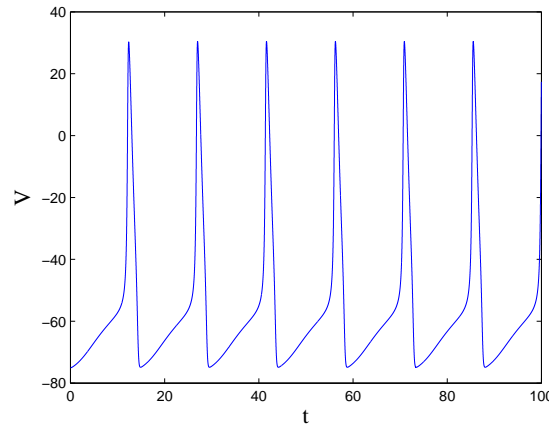


Figure 8.1: The spike train produced by the Hodgkin-Huxley model under constant external current input $I = 10$.

8.2 Morris-Lecar model

For practical reasons, a simpler neuron model based on the work of Morris and Lecar [130] is often considered. This model consists of two differential equations instead of four as in the case of Hodgkin-Huxley model, and its dynamics can therefore be more easily visualised using a planar state-space illustration. The equations of the Morris-Lecar model are:

$$\frac{dV}{dt} = -g_L(V - V_L) - g_{Ca}m_\infty(V)(V - V_{Ca}) - g_Kw(V - V_K) + I, \quad (8.9)$$

$$\frac{dw}{dt} = \lambda(V)(w_\infty(V) - w). \quad (8.10)$$

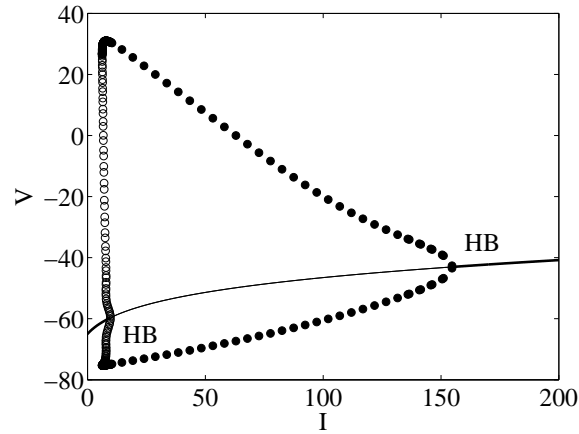


Figure 8.2: Bifurcation diagram of the Hodgkin-Huxley model as a function of the external current input I . Bold lines correspond to stable fixed point, light line to an unstable fixed point. Full (empty) circles correspond to stable (unstable) periodic orbits. Two Hopf bifurcations are denoted as ‘HB’.

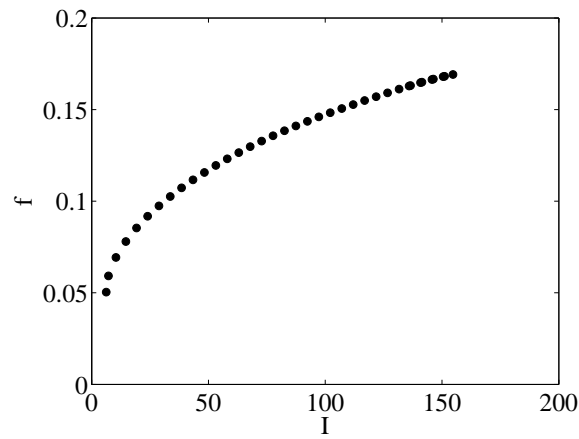


Figure 8.3: Firing frequency of the Hodgkin-Huxley model as a function of the external current input I . Only the frequency for the stable limit cycle is shown (full circles). Note the onset of the oscillations with a non-zero frequency via the Hopf bifurcation. This is characteristic of the so-called *Type II* neural excitability.

The first equation describes the voltage dynamics based on current balance similarly as in the model of Hodgkin and Huxley. The second equation describes the dynamics of the recovery variable w which represents the fraction of open potassium ion channels. Note that the faster calcium dynamics are approximated here to be instantaneous. The voltage dependent gating variables m_∞, w_∞ and voltage dependent rate λ are defined by:

$$m_\infty(V) = \frac{1}{2} \left[1 + \tanh \left(\frac{V - V_1}{V_2} \right) \right], \quad (8.11)$$

$$w_\infty(V) = \frac{1}{2} \left[1 + \tanh \left(\frac{V - V_3}{V_4} \right) \right], \quad (8.12)$$

$$\lambda(V) = f \cosh \left(\frac{V - V_3}{2V_4} \right). \quad (8.13)$$

While originally developed to describe the barnacle giant muscle fiber, the Morris-Lecar model has become one of the favourite conductance-based models in computational neuroscience. Dependent on its parameter settings, it can replicate a wide range of behaviour of various neural cells. When set to exhibit the onset of repetitive firing with arbitrarily low frequency, the model becomes particularly useful for modelling pyramidal neurons of the neocortex. This neural excitability behaviour (sometimes called *Type I*) can be achieved for instance with the following set of parameters: $V_1 = -0.01$, $V_2 = 0.15$, $V_3 = 0.1$, $V_4 = 0.145$, $g_{Ca} = 1$, $g_K = 2.0$, $g_L = 0.5$, $V_K = -0.7$, $V_L = -0.5$, $V_{Ca} = 1$ and $f = 1.15$ as in [77]. This is also the parameter set we use in our analysis and simulation in Chapter 10.

As already mentioned, due to the reduced two-dimensional description, the behaviour of the Morris-Lecar model can be well visualised in the state space of the variables V and w . Each point in this two-dimensional plane uniquely characterises the state of the dynamical system. In the following, we will describe the geometry of the model behaviour. To determine the fixed points of the model, we first find the *nullclines* of the system ((8.9)), i.e. the curves in the state-space that correspond to either $V' = 0$ or $w' = 0$. Depending on the value of the injected current I , the nullclines have one or three intersections, giving rise to one or three fixed points.

The phase-plane portrait of the Morris-Lecar model for the external drive set to $I = 0.076$ is depicted in Figure 8.4. For this value of I , the system has three fixed points – a left-most low-voltage stable one, a middle unstable saddle-node and a stable high-voltage fixed point. There is a small-amplitude unstable limit

cycle and a large-amplitude stable limit cycle around the third fixed point – the latter corresponds to a continuous spike train that can be generated by the neuron under constant current input I .

The behaviour of the neuron can of course change with the level of the current input. This dependence is captured in the bifurcation diagram in Figure 8.5. Here, the most important features of the phase-plane portrait are reproduced as a function of I . The dependence of the firing frequency f on the external input current I is depicted in Figure 8.6. It has the approximate form $f \sim \frac{-1}{\log(I-I_c)}$, where I_c is the critical input current value. We observe an onset of oscillations with zero frequency, characteristic of the *Type I* neural excitability.

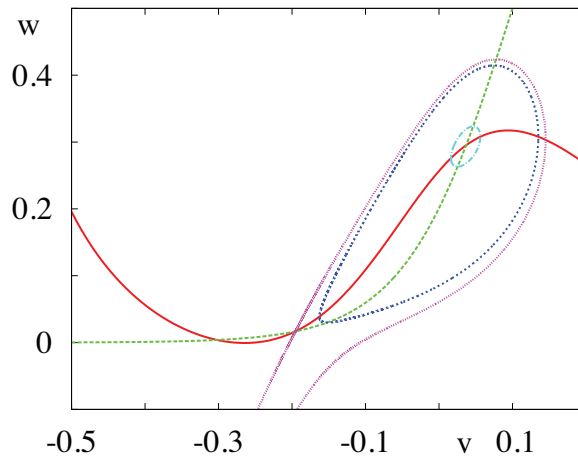


Figure 8.4: State-space portrait of the Morris-Lecar model with the external drive set to $I = 0.076$. The nullclines are shown in green and red. Their intersections define the fixed points. In the order from lower voltage, these are stable, unstable and stable. There is a small-amplitude unstable limit cycle (light blue) and a large-amplitude stable limit cycle (dark blue) around the third fixed point. The separatrix (purple) divides the state-space into the basin of attraction of the leftmost stable fixed point and of the stable limit cycle (and the contained rightmost stable fixed point).

We shall come back to Morris-Lecar model in chapter 10, where we study a synaptically coupled network of Morris-Lecar neurons endowed with regulatory action of endocannabinoids.

8.3 Wilson-Cowan model

An example of a mesoscopic level of description is the Wilson-Cowan model. This is a representative of a class of so-called *population models*, where the activity

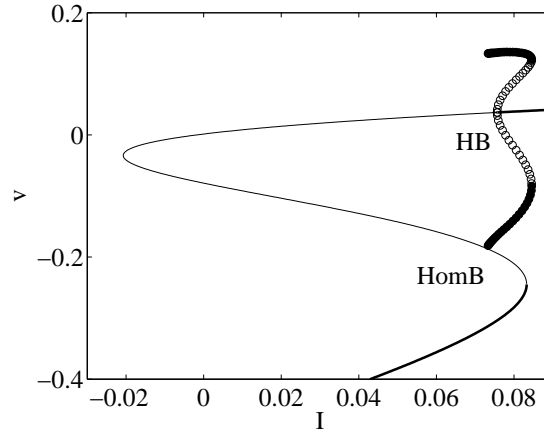


Figure 8.5: Bifurcation diagram of the Morris-Lecar model as a function of the external drive I . The full lines correspond to stable fixed points, the thin lines to unstable fixed points, full (empty) circles denote the maximum and minimum of a stable (unstable) limit cycle. HB denotes a Hopf bifurcation, HomB a homoclinic bifurcation.

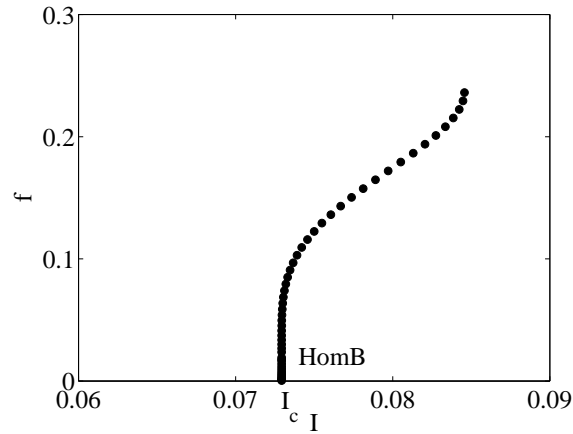


Figure 8.6: Firing frequency of the Morris-Lecar model as a function of the external current input I . Only the frequency for the stable limit cycle is shown (full circles). Note the onset of the oscillations with a zero frequency via the homoclinic bifurcation (HomB). This is characteristic of the so-called *Type I* neural excitability.

of whole populations of neurons is modelled in a simplified form. Importantly, among other applications, this type of models (embedded into anatomically connected networks) has been used previously for modelling the dynamics of large-scale spontaneous brain activity, see e.g. [42]. More detailed investigation of such a population-network model is provided in chapter 9.

The equations of the ‘coarse-grained version’ of the model are:

$$\tau_e \frac{\partial E}{\partial t} = -E + (k_e - r_e E) \mathcal{S}_e(c_1 E - c_2 I + P), \quad (8.14)$$

$$\tau_i \frac{\partial I}{\partial t} = -I + (k_i - r_i I) \mathcal{S}_i(c_3 E - c_4 I + Q). \quad (8.15)$$

These equations describe the evolution of variables E and I which represent the proportion of excitatory (inhibitory) cells firing per unit time at the instant t in a spatially localised neural population. The functions $\mathcal{S}_e, \mathcal{S}_i$ are called the subpopulation *response functions* - they give the expected proportion of neurons receiving at least threshold excitation for given overall input. The general form of these functions is a sigmoidal function - a real-valued monotonous differentiable function $f(x)$ with exactly one inflection point and $\lim_{x \rightarrow -\infty} f(x) = 0$, $\lim_{x \rightarrow \infty} f(x) = 1$. One typical choice of the sigmoid function is the logistic function $f(x) = 1/(1 + e^{-x})$. The argument of the response function \mathcal{S}_e represents the incoming activity of the excitatory subpopulation, where the *connectivity coefficients* c_1, c_2 represent the average number of excitatory and inhibitory synapses per cell and $P(t)$ denotes the external input to the excitatory subpopulation. Similarly c_3, c_4 and $Q(t)$ describe the input to the inhibitory subpopulation. The constants r_e, r_i model the effect of refractoriness of the neurons, implementing the fact that some of the neurons might be in a so-called *refractory period*. In biological neurons, this is a period of time after firing during which the cell is not able to generate a new action potential even when exposed to overthreshold stimulation. Finally, the constants k_i, k_e are included to correct for the fact that the maximal values of the response function are in general not equal to unity.

In the following we describe the behaviour of the Wilson-Cowan model. For our analysis we choose a parameter setting of $k_e = k_i = 1, r_e = r_i = 0, c_1 = 10, c_2 = 10, c_3 = 10, c_4 = -2$ according to setting used in [85]. Importantly, the behaviour of the model still depends on the level of the variables P, Q . These represent baseline constant activity, input from other cells or external experimental stimulation. For example, with the above parameter settings, for

the values $P = Q = 0$ the system rests in its only stable fixed point. On the other hand when we set $Q = -8$, the system oscillates with a period $T \sim 4$ with a stable limit cycle centred around an unstable fixed point.

To characterise the dependence of the qualitative dynamics of the model on the values of the external input parameters P, Q , we carry out a bifurcation analysis. The dependence of the model dynamics on Q is shown in Figure 8.7 for fixed $P = 0$. We see the origin of the oscillations in the Hopf bifurcation at $Q \sim -8.7$ and its disappearance in another Hopf bifurcation at $Q \sim -3.3$. The frequency of the oscillations as a function of Q is depicted in Figure 8.8. On the other hand, fixing the background input of the inhibitory populations at e.g. $Q = -4$, we can explore the dependence of the model behaviour on the parameter P - see Figure 8.9. Note that here the stable limit cycle disappears through a *saddle-node on limit cycle* bifurcation, also known as saddle-node of invariant cycle (SNIC). The frequency dependence is shown in Figure 8.10 - the frequency curve has the approximate form of $f \sim \sqrt{|P - P_c|}$. When the excitatory population drive P is close to the critical drive P_c , we observe oscillations with arbitrarily low frequency. This is a characteristic feature of a *Type I* neural excitability.

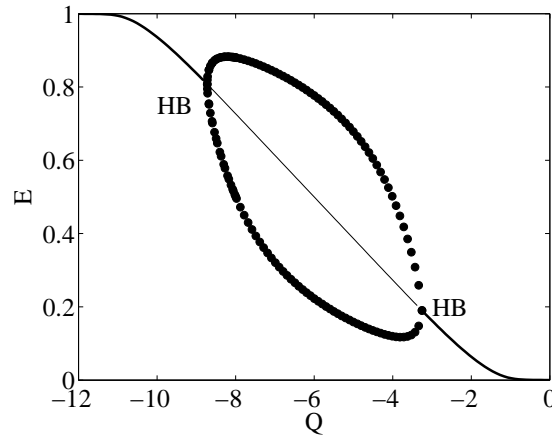


Figure 8.7: Bifurcation diagram of the Wilson-Cowan model as a function of the parameter Q for fixed $P = 0$. The thick lines correspond to stable fixed points, the light line to unstable fixed points, full circles denote the maximum and minimum of a stable limit cycle.

An overview of the dependence of the structure on the two parameters P, Q is given by the two-dimensional bifurcation diagram in Figure 8.11. Some features of the bifurcation diagram can be derived analytically, such as the values of P, Q corresponding to the Hopf bifurcation as shown below.

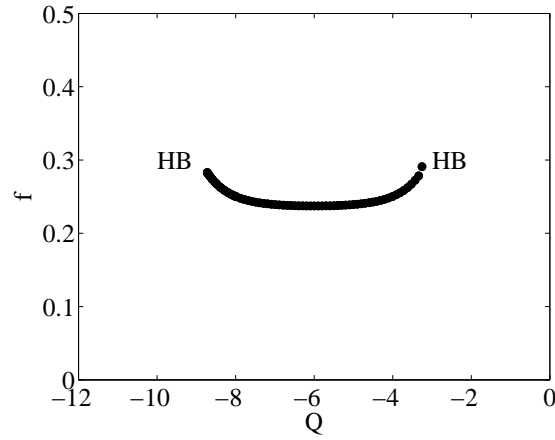


Figure 8.8: Oscillation frequency of the Wilson-Cowan model as a function of the parameter Q for fixed $P = 0$. The full circles correspond to the stable limit cycle.

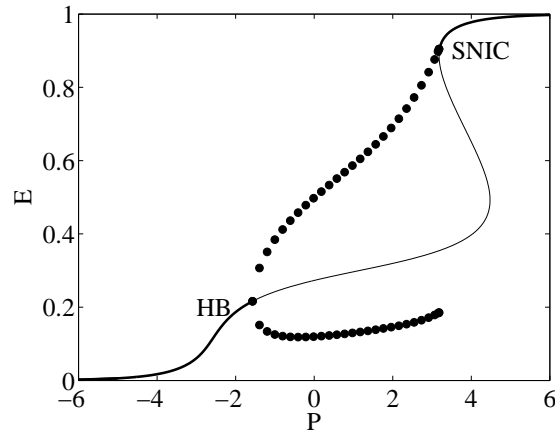


Figure 8.9: Bifurcation diagram of the Wilson-Cowan model as a function of the parameter P for fixed $Q = -4$. The thick lines correspond to stable fixed points, the light line to unstable fixed points, full circles denote the maximum and minimum of a stable limit cycle.

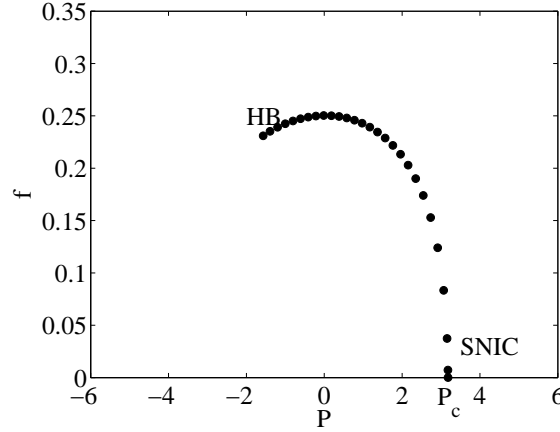


Figure 8.10: Oscillation frequency of the Wilson-Cowan model as a function of the parameter P for $Q = -4$. Note the onset of the oscillations with a zero frequency via the SNIC bifurcation. This is characteristic of the so-called *Type I* neural excitability.

First, for a point (E, I) to be a fixed point, it must hold

$$0 = -E + \mathcal{S}(c_1 E - c_2 I + P), \quad (8.16)$$

$$0 = -I + \mathcal{S}(c_3 E - c_4 I + Q), \quad (8.17)$$

which gives

$$P = \log(E/(1 - E)) - c_1 E + c_2 I, \quad (8.18)$$

$$Q = \log(I/(1 - I)) - c_3 E + c_4 I. \quad (8.19)$$

Therefore the Jacobian matrix at the fixed point is

$$L = \begin{vmatrix} -1 + c_1 E(1 - E) & -c_2 E(1 - E) \\ c_3 I(1 - I) & -1 - c_4 I(1 - I) \end{vmatrix}. \quad (8.20)$$

We can now derive the position of the Hopf bifurcation using the conditions:

$$\text{Tr} L = -2 + c_1 E(1 - E) - c_4 I(1 - I) = 0 \text{ and } \det L > 0. \quad (8.21)$$

The former equality allows us to eliminate I as

$$I_{\pm}(E) = \frac{1 \pm \sqrt{1 - 4(-2 + c_1 E(1 - E))/(-c_4)}}{2}, \quad (8.22)$$

while the latter condition further limits the location of the Hopf bifurcation. Now

we can plot the Hopf bifurcation fixed point equations (8.18) parametrically in the $P(E), Q(E)$ plane as shown in Figure 8.11.

The same approach can be used to derive the position of the saddle-node bifurcation in the $P(E), Q(E)$ plane. The presence of the bifurcations in Figure 8.11 was further confirmed by numerical simulations and bifurcation continuations using XPPAUT[46] software.

Most importantly, for the choice of parameters in the central area between the two Hopf and two SNIC bifurcations, the Wilson-Cowan model oscillates with a stable limit cycle. On the other hand, when this boundary is crossed the system goes to a stable fixed point.

Nevertheless, the diagram shows only the main skeleton of the bifurcation structure of the system. In particular, the triangular areas between the lines denoting HB, SN and SNIC bifurcation contain a relatively rich structure of bifurcations (see [85] for more detail).

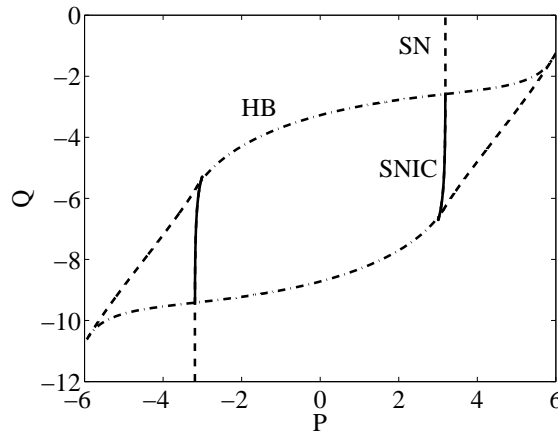


Figure 8.11: Two-dimensional bifurcation diagram of the Wilson-Cowan model. Full lines denote location of saddle-node on limit cycle bifurcations (SNIC), dashed lines denote simple saddle-node bifurcations (SN), dash-dotted lines denote Hopf bifurcations (HB).

8.4 Networks of weakly coupled oscillators

In this section we briefly introduce some concepts that are useful for description and analysis of networks of neural oscillators, particularly in the case when they are interlinked by weak interactions. This theory is going to be used later in chapter 10 for the analysis of a network of weakly coupled Morris-Lecar neurons.

8.4.1 Introduction

So far we have dealt with models of a single neuron or a population, without explicitly modelling the synaptic interactions. Now we will consider a system of n identical neuronal oscillators, each with dynamics described by

$$\dot{X}_i = F(X_i), \quad X_i \in \mathbb{R}^m, \quad i = 1, \dots, n, \quad (8.23)$$

for which the coupled system is given by

$$\dot{X}_i = F(X_i) + \epsilon G(X), \quad i = 1, \dots, n, \quad (8.24)$$

where G denotes the coupling function. In this section we describe under what conditions and how we may reduce a general system of neural oscillators to a phase model of a form

$$\dot{\theta}_i = \Omega_i + \epsilon g_i(\theta_1, \dots, \theta_n), \quad \theta_i \in \mathbb{S}^1, \quad i = 1, \dots, n \quad (8.25)$$

defined on the n -torus $\mathbb{T} = \mathbb{S}^1 \times \dots \times \mathbb{S}^1$. This effectively offers a simplified description of the original network dynamics. Subsequently we introduce basic techniques for determining stability of phase-locked solutions of such a phase model.

8.4.2 Phase reduction

The reduction is given by the following theorem from [85]:

Consider a family of weakly connected systems given by (8.24) such that each equation in the uncoupled system ($\epsilon = 0$) given by (8.23) has an exponentially stable limit cycle attractor $\gamma_i \subset \mathbb{R}^m$ having natural frequency $\Omega_i \neq 0$. Then, the dynamical system (8.25) is a local model for (8.24). That is, there is a neighbourhood W of $M = \gamma_1 \times \dots \times \gamma_n \subset \mathbb{R}^{mn}$ and a continuous function $h : W \rightarrow \mathbb{T}^n$ that maps solutions of (8.24) to those of (8.25).

The proof of this theorem is based on the fact that the product of the exponentially orbitally stable limit cycles M is a normally hyperbolic invariant manifold. This further guarantees the existence of a continuous function mapping the solutions of the system (8.24) to solutions of a corresponding system defined directly on the invariant manifold M . For details of this part of the proof we refer the reader to [85]. In the second part of the proof, a continuous mapping between

the invariant manifold M and \mathbb{T}^n is found, that maps the solutions of the respective systems. This part is outlined below in detail as it demonstrates the relation between the original and the reduced model.

Since γ_i is homeomorphic to the unit cycle \mathbb{S}^1 , we can parametrize it using the natural phase variable $\theta_i \in \mathbb{S}^1$; $\theta_i = \Omega_i t$. The natural parametrization is given by the mapping

$$\Gamma_i : \mathbb{S}^1 \rightarrow \gamma_i, \quad \Gamma_i(\theta_i(t)) = X_i(t) \in \gamma_i, \quad t \in [0, 2\pi/\Omega_i], \quad (8.26)$$

with an arbitrary but fixed choice of origin $X_i(0)$ of the cycle and $X_i(t)$ being the periodic solution to $\dot{X}_i = F_i(X_i)$. Then we have:

$$\dot{X}_i = \frac{d\Gamma_i(\theta_i(t))}{dt} = \Gamma'_i(\theta_i)\dot{\theta}_i = \Gamma'_i(\theta_i)\Omega_i = F_i(\Gamma_i(\theta_i(t))) \quad (8.27)$$

for all t . This further gives,

$$\Gamma'_i(\theta_i) = F_i(\Gamma_i(\theta_i))/\Omega_i \quad (8.28)$$

for all θ_i in \mathbb{S}^1 . When we substitute $X_i(t) = \Gamma_i(\theta_i(t))$ from (8.26) and $\dot{X}_i = \Gamma'_i(\theta_i)\dot{\theta}_i$ from (8.27) into (8.24), we obtain:

$$\Gamma'_i(\theta_i)\dot{\theta}_i = F_i(\Gamma_i(\theta_i)) + \epsilon G_i(\Gamma(\theta)), \quad (8.29)$$

where $\theta = (\theta_1, \dots, \theta_n)$. We further modify the left side using (8.28) and subsequently multiply both sides of the equation by

$$R_i(\theta_i) = \frac{\Omega_i F_i(\Gamma_i(\theta_i))^T}{|F_i(\Gamma_i(\theta_i))|^2} \quad (8.30)$$

to obtain

$$\dot{\theta}_i = \Omega_i + \epsilon R_i(\theta_i) G_i(\Gamma(\theta)). \quad (8.31)$$

This equation already has the form of (8.25). We shall call the function $R(\theta)$ *phase response curve* (PRC), as it describes the dependence of the perturbation response of the phase of the oscillator on the current phase of the oscillator. Note that $R(\theta)$ is defined on the torus \mathbb{S}^1 and therefore periodic. The initial value of the PRC depends on the choice of the starting point of the natural parametrization of the limit cycle of the original oscillator.

While the PRC summarizes the information about the behaviour of a single oscillator of the network with respect to its phase perturbation, to determine the stability of particular network solutions, one needs to take into account the properties of the connections as well. In the following we consider a case of simple synaptic coupling. Let the input of the i -th oscillator in the system be:

$$G_i(t) = \sum_{j=1}^N W_{ij} s_j(t), \quad s_j(t) = \sum_m \eta(t - T_j^m), \quad (8.32)$$

where $s_j(t)$ represents the synaptic activity of the j -th neuron, T_j^m is the m -th firing time of the j -th neuron and W_{ij} represents the anatomical connectivity from the j -th to the i -th oscillator. The synaptic function $\eta(t)$ defines the stereotypical temporal profile of synapse activity in response to action potential generation.

We further consider a case when all the oscillators are identical, so that we can drop the index i for the phase response curve $R(\theta)$ and for the natural frequency Ω . For convenience, rotating coordinates ψ_i can be also introduced by $\theta_i = \psi_i + \theta_i \Omega$.

As for the firing times it holds that $\theta_j(T_j^m) = m, m \in \mathbb{Z}, j \in 1, \dots, N$, for the firing times we have $T_j^m = (m - \psi_j)\Delta$. Substituting into the synaptic activity we have $s_j(t) = \sum_m \eta(t - T_j^m) = \sum_m \eta((\theta_i - m)\Delta)$. Substituting into (8.25) we obtain an equivalent system of ODEs for the new variables ψ_i :

$$\dot{\psi}_i = \epsilon \sum_j W_{ij} R(\psi_i + \Omega t) P(\psi_j + \Omega t), \quad (8.33)$$

where

$$P(\theta) = \sum_{m \in \mathbb{Z}} \eta((\theta + m)\Delta), \quad 0 \leq \theta < 1, \quad \eta(t) = 0, t < 0. \quad (8.34)$$

The phase evolution equations can be further simplified using the averaging approximation, where the input from the j -th to the i -th oscillator is averaged over the natural period Δ . Therefore we obtain:

$$\begin{aligned} \dot{\psi}_i &= \epsilon \sum_j W_{ij} \frac{1}{\Delta} \int_{-\psi_j \Delta}^{-\psi_j \Delta + \Delta} R(\psi_i + \Omega t) P(\psi_j + \Omega t) dt = \\ &= \epsilon \sum_j W_{ij} H_{syn}(\psi_j - \psi_i), \end{aligned} \quad (8.35)$$

where

$$H(\psi) = \int_0^1 R(\theta - \psi) P(\theta) d\theta \quad (8.36)$$

is called *phase interaction function*.

8.4.3 Stability of phase locked solutions

An important type of a synchronous solution of a system of oscillator equations is a so-called 1:1 *phase locked solution* of a form $\theta_i = \phi_i + t/T$, where ϕ_i is a constant phase offset of the i -th oscillator and T is the collective period of the coupled oscillators.

The stability of such a phase locked solution can be investigated by the following standard approach. Consider a perturbation $\tilde{\theta}(t)$ of the solution $\theta(t)$ in a vicinity of the phase locked solution. The perturbation of each oscillator is defined by

$$\tilde{\theta}_i(t) = \theta_i(t) - \phi_i - \Omega t. \quad (8.37)$$

Substituting to the equation (8.35) we obtain a system of equations for the perturbation:

$$\frac{d\tilde{\theta}_i}{dt} = \epsilon \sum_j W_{ij} H((\phi_j - \phi_i) + (\tilde{\theta}_j - \tilde{\theta}_i)). \quad (8.38)$$

To determine stability of the phase locked solution, we linearise the perturbation equation to the form

$$\frac{d\tilde{\theta}_i}{dt} = \epsilon \sum_j \hat{H}_{ij}(\Phi) \tilde{\theta}_j, \quad (8.39)$$

where $\epsilon \hat{H}_{ij}$ is the Jacobian matrix corresponding to the equation of the perturbation system (8.38), i.e.

$$\hat{H}_{ij} = W_{ij} H'(\phi_j - \phi_i) - \delta_{i,j} \sum_k W_{ik} H'(\phi_k - \phi_j), \quad (8.40)$$

and $H'(\phi)$ denotes the derivative of $H(\phi)$ with respect to the argument ϕ .

The eigenvalues of the Jacobian \hat{H} now determine the stability of the phase locked solutions of the original system. In particular, one of the eigenvalues is always zero, with the corresponding eigenvector $(1, 1, \dots, 1)$ pointing in the direction of the flow of the phase locked solution. The other values can take in the general case complex values. However, the phase locked solution is stable under the condition that all the other eigenvalues have negative real parts [47], i.e. perturbations in all directions apart from the direction of the flow asymptotically diminish.

8.5 Models of spontaneous brain activity

Over the last decade the interest of the neuroimaging community in spontaneous brain activity has steadily increased. Recently this trend has attracted the attention of the computational modelling community to this topic, leading to the first attempts to capture the large-scale spontaneous spatiotemporal dynamics observed in resting state fMRI in mathematical models.

Less than three years ago a key paper [82] attempted to describe the relationship between the underlying anatomical connectivity and the fMRI BOLD functional connectivity pattern using a dynamical neural model. This consisted of 47 neural oscillators representing a selection of cortical areas within a macaque brain hemisphere, linked by structural connections informed by macaque anatomical connectivity studies collated in the neuroinformatics tool CoCoMac [102]. An important outcome of this study was the observation of a strong correlation between the input structural connectivity matrix and the functional connectivity matrix computed from the time series produced by the model. While the details of the model used and the interpretation of quite complex analysis of the results can be criticized from several perspectives, this study served as a proof of principle that explicit modelling of spontaneous large-scale brain activity is feasible. Moreover, this made the wider community aware that, at least in principle, such modelling studies can provide predictions that might be of interest for experimentalists and conversely lead to amendments in the model in case of disagreement with the observed empirical results. In the following two years, several new studies using this or alternative models have been published in the area.

To start with, the model of the pioneering paper was so far used in three other related works [83, 84, 4]; the last two articles already use a human structural

connectivity matrix obtained using tractography applied on Diffusion Spectrum Imaging MRI data. Two of these papers focus on using the model to explore potential changes in functional connectivity under lesions of the underlying structural substrate of the model. While such an application would be of immense interest if reasonably accurate, at the moment the predictions are rather general and were not directly validated on real data.

While the model discussed above is based on deterministic, although most likely chaotic, dynamics, a different model has been implemented in other studies [42, 63]. The authors here implement noise and explicit signal transmission delays in their models and argue that these have a key role in the emergence of spontaneous brain fluctuations.

For completeness it is important to note that the above models, directly motivated by the observed functional connectivity in resting state fMRI, seem to have developed relatively independently of a somewhat older stream of research into the relation between brain anatomy and brain functional organization, that particularly focused on brain modularity. A representative recent example of such work is [93] providing further references to older works in the area. Models of this group typically use very simplified models of neurons and their interaction, falling into the category of *artificial* rather than *biological neuronal networks* modelling.

It is particularly worth mentioning the work of the Potsdam group for instance the concept of *networks of networks* described e.g. in [181] is a promising approach for modelling multiscale neural dynamics. Here, the connectivity is modelled on two levels - the local connectivity patterns are modelled using localised small-world networks, while the global connectivity pattern is defined by a cortico-cortical anatomical connectivity matrix derived from histology.

8.5.1 Model architectures

The models proposed in recent years for spontaneous brain activity share some key features. In particular, typically several dozens of cortical areas are included in the model, whilst subcortical areas are usually omitted. The structural connections between the areas are also usually defined by a matrix derived from some anatomical study. One option used is a macaque monkey anatomy collated in the CoCoMac database [102], where versions with 47 (as in [82]) or 38 (e.g. in [63]) nodes are typically used. More recently, reasonably reliable human

whole head anatomical connectivity studies using tractography based on Diffusion Spectrum Imaging (DSI) have become feasible, which enabled the use of human connectivity matrices in modelling studies such as in [84]. Importantly, high-quality acquisition and analysis is still a technically demanding task, with the methods under constant development. Also, the already acquired matrices were not yet made publicly available, limiting the use to direct collaborations.

On the other hand, details of the studies differ significantly. While each of the cortical areas is always represented by a single neural oscillator, the choice of the oscillator type varies widely across the modelling studies. Honey et al. [82, 83, 84, 4] use a hybrid model based on the Morris-Lecar single neuron model. Ghosh et al. [63] use a FitzHugh-Nagumo model, although they state that they obtained ‘similar’ results with other models. As they study the behaviour of the model always in a close vicinity of a Hopf bifurcation and tune the parameters, this might be caused by a local equivalence of the models - unfortunately they do not provide details of the implementation or results for the other models. Finally Deco et al. [42] have used a Wilson-Cowan *population* model at each cortical node.

The studies also differ in how to arrive at sustained activity of the model. While Honey et al. [4, 82, 83, 84] set the parameters such that the single modules are in the oscillatory regime, other researchers typically position the single modules in the regime with a stable fixed point, but close to instability due to a Hopf bifurcation. This system is then driven by a white noise, which leads to transient oscillations of the system. The latter models also include cortico-cortical delays in the model, and it is argued that the noise and delays play a crucial role in the emergence of the typical spontaneous brain activity features [42, 63].

8.5.2 Explanatory potential

The ultimate goal of modelling spontaneous brain activity might be seen e.g. in explaining alterations of resting state activity seen in changed brain state or in patients with a wide range of diagnoses (see chapter 3 for more details of results of neuroimaging studies). Nevertheless, the initial step is to be able to elucidate the emergence of the two mutually interlinked characteristic features observed in resting state data of healthy human subjects - functional connectivity and low-frequency fluctuations. These features are actually quite stable across conditions, disease states and even across species.

The level to which this modelling goal has been accomplished by the published studies is questionable. While some promising results have been presented that form a proof of principle that such an enterprise is realistic, a closer observation reveals that much of the evidence provided is indirect or potentially confusing.

When it comes to explaining the functional connectivity patterns, the initial study of Honey et al. [82] presented a promising result of up to 80% overlap of the structural connectivity (SC) matrix with the thresholded model-based functional connectivity matrix (the details of the overlap measures are reviewed in the following section). Nevertheless, a later study by the same group [84] argues that thresholding of resting state FC yields highly inaccurate prediction of SC – both for FC matrices obtained from real data and model simulations.

Even more disappointing seems to be the point not particularly highlighted but documented in the paper, that the structural connectivity itself was in many cases a better predictor of the real functional connectivity than the simulated functional connectivity computed using the model. Also, in many aspects a linearised version of the model outperformed the nonlinear version, as documented in the Supplementary Information of the paper.

The most promising reproduction of functional connectivity patterns so far has been presented in the study by Ghosh et al. [63]. At realistic levels of critical signal transmission velocity, they reported the model to be able to predict correctly the status (presence or absence) of 14 out of 15 functional connections as presented in one of the original region-wise FC resting state studies [51]. This is an interesting result which calls for independent validation on a different dataset, including more than the presented 6 regions and avoiding the comparison across species. A unique feature of this study is that it claims that these results are present only for a narrow window of parameters of the model – a point that would be worth a closer investigation.

Finally, the study by Deco et al. [42] proposes to explain anti-correlation of within-cluster synchronisation of strength $r = -0.1$ at a narrow window of parameters corresponding to typical frequency of 0.1Hz by a mechanism of stochastic resonance. This raises a number of questions. First, although anti-correlations between time series of particular resting state networks have been reported, the extent of their validity is still a matter of a heated debate in the neuroimaging community and therefore they might not form the best feature for assessing the models against (more detail on the problem is covered in subsection 4.1.2). Second, this study reports anti-correlations in cluster global

synchronisation, rather than local signal levels as reported in the neuroimaging studies. Although the two might be linked in particular cases, such a comparison is at the very minimum confusing. Finally, the frequency of 0.1Hz is *not* very typical of resting state LFF, which are commonly almost an order of magnitude slower.

This leads us to the question of how successful the recent studies have been in modelling the second cardinal feature of spontaneous brain activity, i.e. the LFF [36] (typically 0.01 – 0.1Hz).

To start with, we have to deal with a basic, although often overlooked, property of the BOLD signal. Namely, acquired BOLD MRI signal is a result of relatively sparse sampling of a haemodynamic response to increased demands of the neural tissue. This response is not immediate, but effectively acts as a (non-linear) low-frequency filter. The haemodynamic response has a typical gradual onset starting 2 seconds after the stimulus, rising to a plateau at 6-9 seconds after the stimulus onset and then slowly returns to the baseline [113].

From this perspective, the prevalence of slow components in the frequency spectra of real or modelled BOLD signal is rather trivial than surprising. Indeed, a low-frequency signal can be produced even from white noise by simple filtering with a linear approximation of the haemodynamic response function. Nevertheless, it has been shown that electrophysiological signals that are free of the neurovascular low-frequency filters *also* show LFF in the band-limited power [110]. Further, the band-limited power fluctuations of EEG have been repeatedly shown to correlate with the local BOLD signal (see e.g. [108, 118]). Also, band-limited powers of direct electrocorticographic measurements in human patients have shown correlation and reactivity patterns similar to the BOLD-derived resting state networks [127, 135]. While due to technical reasons the electrophysiological studies of LFF are severely limited, there seems to be sufficient evidence for LFF in local activity independent of the smoothing action of the haemodynamic response. Therefore it seems reasonable to require that the models of spontaneous brain activity generate LFF not only on the level of the BOLD signal.

Taking this into account, the models presented above seem to generally fall short of explaining the experimental LFF observations. Typically, they do mention LFF property of some version of the signal, but the relation and relevance for the signal observed in neuroimaging is often vague. For example, in [82] the slow dynamics is discussed for the global synchronisation measured by mean transfer entropy (TE, [157]) between the time series. While this finding does

not have a direct correlate in available neuroimaging studies, the result of slow dynamics is not completely surprising given that the TE was computed for 30-second sliding windows with 6 seconds sampling rate, corresponding to 24-second overlaps. We have also already commented on the weak points of the finding of ~ 0.1 Hz fluctuations in between-module synchronisation difference [42]. On the other hand, the finding of low-frequencies in power spectra of simulated BOLD signal in [63] may be a trivial result of the haemodynamic response modelling. Unfortunately, the reviewed modelling studies do not show explicitly nontrivial low-frequency content in signal types that would be directly comparable to the electrophysiological correlates of the BOLD LFF. Although some of the models might produce such comparable behaviour, in the absence of direct published evidence their ability to explain the LFF feature remains a matter of speculation.

In summary, we can characterize the field of modelling spontaneous brain activity features observed in neuroimaging as being in an initial stage. It offers promising concepts but provides only rather general if any explanations and predictions, especially compared to the wealth of data gathered during the recent decade by the experimentalists. The game of communication and integration of these two research aspects in this field is in the early stages with the main responsibility for further progress currently on the side of the modellers. In the following chapters, we dive into the challenges of spontaneous brain activity modelling in more detail. Particular attention is paid to the relationship between structural and functional connectivity and even more so to candidate mechanisms for generation of the underlying LFF.

Exploration of a large-scale interaction model

In this chapter we further discuss some of the issues that arise in modelling spontaneous brain activity with models such as those reviewed in the previous chapter. We are particularly interested in the potential of similar models to explain the main features of resting state brain activity as observed with neuroimaging methods. These are prominently the low-frequency fluctuation and the functional connectivity patterns introduced in Chapter 3 and analysed in some detail in Part I of this Thesis.

9.1 Introduction

To guide our discussion we use a simplistic model of large-scale spontaneous brain activity to illustrate some pertinent points regarding spontaneous brain activity modelling efforts. The model is similar to those detailed in Section 8.5 in that it in principle divides the cortex into several tens of functional areas, identifies each with a simple neural oscillator and uses an anatomical connectivity matrix of connections to describe the links between the models. In agreement with e.g. [4, 82, 84], but in contrary to e.g. [42, 63], we do not embed any explicit noise or delays in the model. Also, in contrast to the somewhat confusing choice of single neuron models for each node representing a cortical area such as in [63], we use a proper population model. This latter strategy was also adopted e.g. in [42].

As already discussed, *functional connectivity* is a central concept in the study of complex brain dynamics, defined as *temporal correlations between remote neu-*

rophysiological events [56]. The functional connectivity study covers the aspect of functional integration of brain function as opposed to the functional segregation/localisation aspect. One of the aims of modelling (spontaneous) brain activity is to explain the patterns of temporal synchronisation of activity within brain networks, as captured by functional connectivity analysis of neuroimaging data. The role of structural connectivity as the underlying substrate for the functional connectivity has been repeatedly suggested.

By exploring the behaviour of a simple model we can show some important challenges of the spontaneous brain activity modelling enterprise. Most importantly, we illustrate two points. Firstly, the functional connectivity matrix, i.e. the dependence structure of the resulting time series, crucially depends on the choice of parameters of the model. Secondly, we discuss the challenges in modelling the low frequency fluctuations and introduce a simple mechanism for emergence of LFF in population networks that does not rely on delays or noise.

9.2 Model description

In the following we study a large-scale cortical model consisting of interconnected neural populations. For simplification, we assume that a parcellation of the cerebral cortex into functional areas is available, such that each area corresponds to a functional unit that can be represented by a single instance of a localised neural population model. We also assume that the approximate structure of neural connections between the functional areas is available in the form of a connectivity matrix. The limitations of these and other assumptions of this type of model are discussed later.

We choose here the parcellation of the cerebral cortex and the structural connectivity matrix in agreement with [82] - 47 areas of macaque cortex together with an anatomical connectivity matrix collated in the CoComac database [102] are used. Note that for the purposes of this chapter the specific choice of cortical parcellation scheme and structural connectivity is not crucial as we shall comment rather on the general properties of such a type of model. Each cortical area in the model is represented by a Wilson-Cowan module – a model of two interacting populations of neurons – the excitatory and inhibitory neuron pool. The Wilson-Cowan model was described in detail in Section 8.3. Using the parameter setting $k_e = k_i = 1, r_e = r_i = 0, c_1 = 10, c_2 = 10, c_3 = 10, c_4 = -2$ in agreement with [85] it gives the following system of two ODEs for the activity

levels E, I of the two populations:

$$\tau \frac{\partial E}{\partial t} = -E + \mathcal{S}(10E - 10I + P), \quad (9.1)$$

$$\tau \frac{\partial I}{\partial t} = -I + \mathcal{S}(10E + 2I + Q), \quad (9.2)$$

where $\mathcal{S}(x) = 1/(1 + e^{-x})$ is a sigmoid response function. In general, depending on the choice of the non-specific inputs P, Q to the two populations, the Wilson-Cowan model either exhibits oscillation due to a stable limit cycle or decays to a stable fixed point (which is the only stable attractor in the other parameter regimes). For simplicity, in the further description, we set the value of the temporal parameters to $\tau_e = \tau_i = \tau = 1$. Note that changing the common temporal parameter τ corresponds to a simple change in the time-scale of the model. Most commonly the time-scale is considered to be in milliseconds, but by changing the τ parameter it can be reset arbitrarily - a common convenient choice is such that the oscillatory behaviour corresponds to that of typical oscillations in the neuronal populations of interest, e.g. the cortical 10Hz alpha rhythm.

Starting from this ‘cortical module’ model as a building unit, we now construct a large-scale network by connecting 47 of these units by mutual interactions via input to the excitatory populations. In particular, we write:

$$\frac{\partial E_i}{\partial t} = -E_i + \mathcal{S}(10E_i - 10I_i + P + \gamma \sum_{j \neq i} w_{i,j} E_j), \quad (9.3)$$

$$\frac{\partial I_i}{\partial t} = -I_i + \mathcal{S}(10E_i + 2I_i + Q), \quad (9.4)$$

where E_i, I_i denote the activity of the excitatory and inhibitory populations at the i -th cortical area and $W = (w_{i,j})_{i,j \in \{1, \dots, N\}}$ is the matrix of inter-regional connection strengths between the $N = 47$ regions with zeroes on the diagonal, $w_{i,j}$ being the connection strength from the j -th to the i -th area. The binary connectivity matrix used here is shown in Figure 9.1. The parameter γ is a global connectivity scaling parameter - we use various values between 0 and 1 in the analysis below. The model is simulated in MATLAB using the ‘ode45’ solver – an implementation of the Runge-Kutta algorithm with an adaptive step. For each run, the model was sampled at t ranging from 0 to 10000 in $\Delta t = 0.1$ steps. For functional connectivity analysis, the first 1000 steps were discarded to allow for initial transients. An example of the system behaviour is shown in Figure 9.2. All the local populations show oscillations, with slight differences in

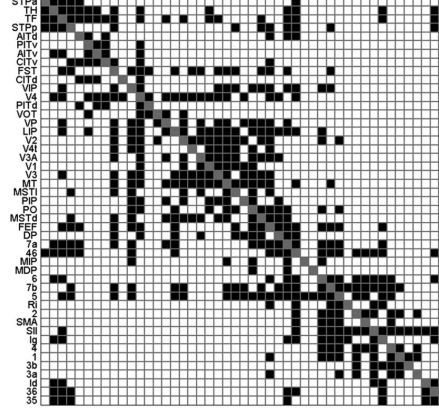


Figure 9.1: A connectivity matrix of 47 regions in the macaque hemisphere. Redrawn from [82].

amplitude and frequency.

9.3 Relation between structural and functional connectivity

In this section, we use our cortical model to show how the functional connectivity changes even when the same structural connectivity matrix and model is used and only the model parameters are varied. We document this by computing the overlap between the underlying structural and emerging functional connectivity matrix similarly as in [82].

9.3.1 Analysis details

As the original structural connectivity matrix is binary with a fixed number of entries, to compute this overlap we first binarise the functional connectivity matrix by thresholding - the threshold is chosen such that the number of over-threshold entries is equal to the number of non-zero entries in the SC matrix. The overlap is then computed simply as the ratio of the region-pairs connected in *both* SC and FC matrices to the total number of connections present in the SC matrix. This index has been used in [82], where they have obtained overlap of up to 0.79, arguing that the aggregate strength of functional couplings between regions is, on average, a good indicator of the presence of an underlying structural link. In their analysis, the maximum value was obtained for TE used as a FC

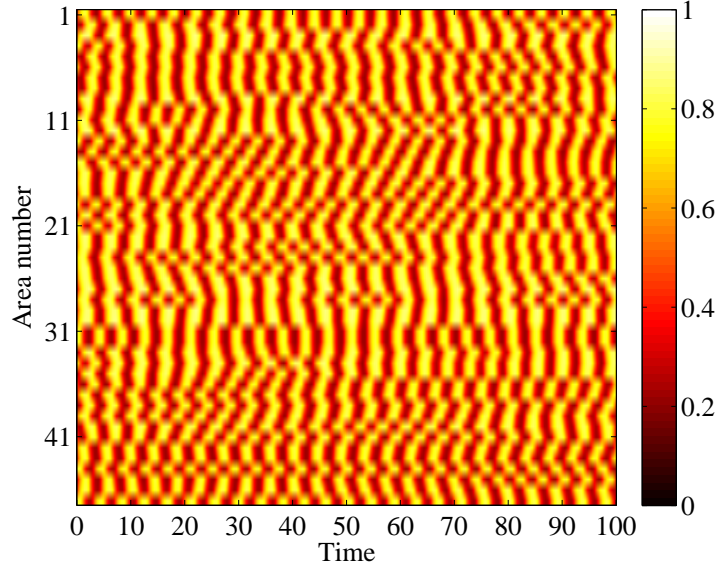


Figure 9.2: An example behaviour of the Wilson-Cowan network. Parameters set to $P = 0, Q = -6, \gamma = 0.1$. The intensity denotes the variable E_i , activity of the excitatory subpopulation of each Wilson-Cowan module. 1000 timesteps are shown from the end of the simulation run (to avoid initial transients).

measure and for very long runs (240,000 timepoints) of their model. Notably, in a later study by the same group [84] the authors concluded somewhat in contrary that thresholding of resting state FC yields highly inaccurate prediction of SC, as discussed in section 8.5.

In our analysis, we measure functional connectivity primarily by linear correlation for simplicity, but include some results for TE later for comparison. Before functional connectivity for each pair of nodes was computed by linear correlation, each univariate time series was monotonically resampled to approximately adhere to gaussian distribution similarly as in [82]. This resampling involved assigning rank to each value in the time series according to their magnitude and subsequently replacing the values by values with the same rank in a sample of appropriate cardinality drawn from the standard normal distribution. The resulting functional connectivity matrix was then thresholded as described in the previous paragraph and the overlap with the structural connectivity matrix was computed.

This process was repeated for a wide range of parameter values P, Q varying in small steps of 0.125. Thus, the graph of the dependence of the overlap on the model parameters is generated - see Figure 9.3. When the pa-

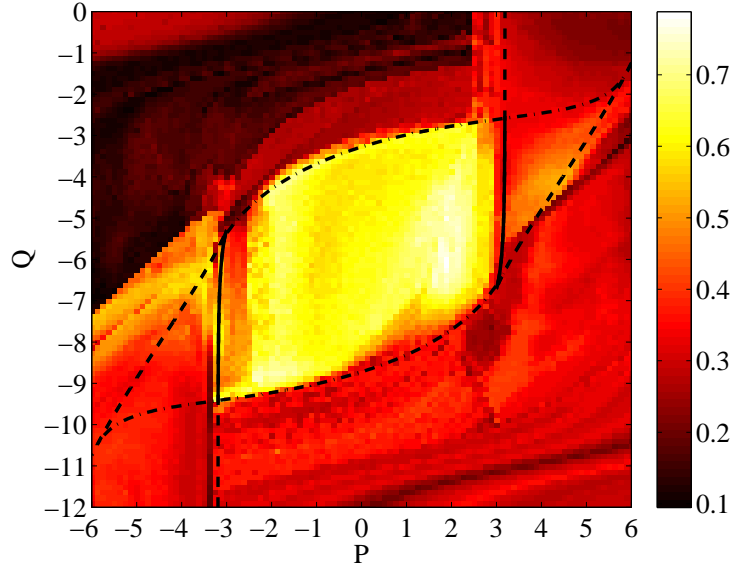


Figure 9.3: The agreement between structural and functional connectivity as a function of the Wilson-Cowan network model parameters P, Q . The colourcoding represents the ratio of the entries shared in the two binary connectivity matrices. Overlaid is the bifurcation structure of a single Wilson-Cowan neural oscillator for orientation - see section 8.3 for details. The central region bordered by the Hopf and SNIC bifurcations corresponds to full-blown oscillations in a single oscillator model. The parameter γ was set to 0.1 for this set of simulations.

rameters P, Q are set such that a single Wilson-Cowan module is in the oscillatory regime, the overlap between the observed functional connectivity and the underlying structural connectivity is clearly above the level expected by chance. Note that for our structural connectivity matrix this ‘chance’ overlap is $N_{connections}/(N_{nodes} * (N_{nodes} - 1)) = 505/2162 \sim 0.23$. Nevertheless, the computed overlap varies widely in an approximate range of 0.2 to 0.8 within the ‘oscillatory’ parameter settings.

This documents the extent to which the correlation structure may depend on other parameters of the system apart from the matrix of relative interconnection strengths. This observation was confirmed by another set of simulations when the overall connectivity has been weakened by setting $\gamma = 0.01$ in equation (9.3). The range of overlaps for different parameter settings remains quite broad. Nonetheless, the overlap values and even the form of dependence on the parameters P, Q was noticeably altered, as shown in Figure 9.4. We hypothesise that for other models a similar range in SC/FC agreements could be observed when varying model parameters.

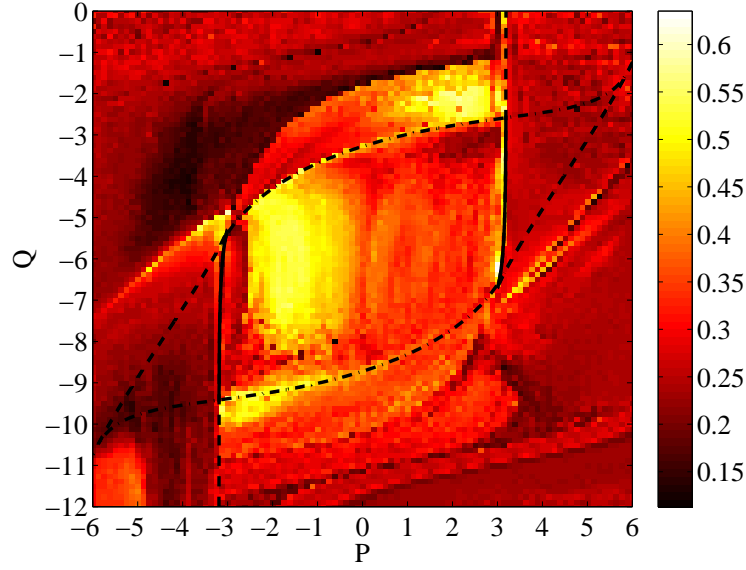


Figure 9.4: The agreement between structural and functional connectivity as a function of the Wilson-Cowan network model parameters P, Q . Graphical display as in Figure 9.3. The parameter γ was set to 0.01 for this set of simulations.

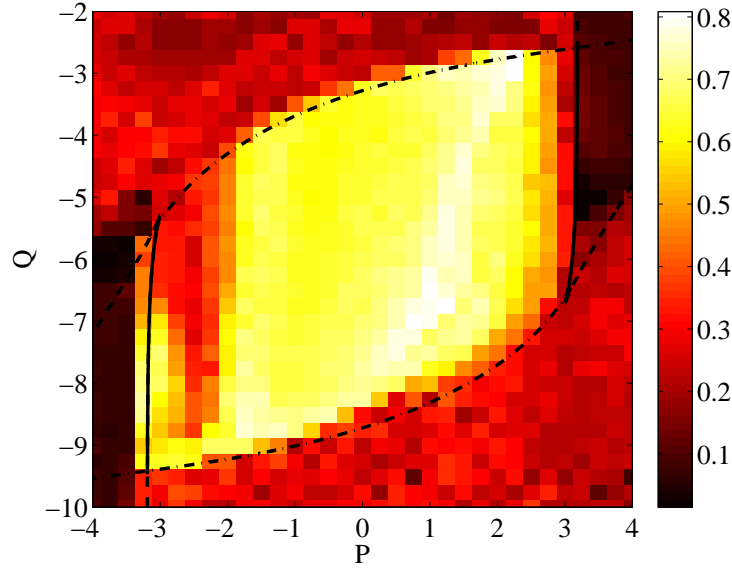


Figure 9.5: The agreement between structural and effective connectivity (measured by transfer entropy) as a function of the Wilson-Cowan network model parameters P, Q . Graphical display as in Figure 9.3. The parameter γ was set to 0.1 for this set of simulations.

9.3.2 Discussion

The observation of mismatch between the structural and functional connectivity is in itself not counterintuitive. There is no theoretical argument for a complete agreement between structural and functional connectivity in the sense that correlation matrix of any dynamical system should equal to that of the matrix of coupling coefficients. What deserves more attention is the fact that the level of agreement can be quite low for some parameter settings. This is in contrast with the view that the functional and structural connectivity closely match, which is to various levels proposed in the recent literature [82, 175]. There are several reasons for the mismatch. In the following we generally assume that functional connectivity is computed by linear correlation of the time series, although the argument remains valid to variable level even for other functional connectivity measures.

First, the functional connectivity matrix is inherently symmetric, which is unlikely to be the case for the underlying structural connectivity matrix. This effectively puts an upper bound on the overlap, with its stringency critically dependent on the level of asymmetry of the structural connectivity matrix. In the case of the matrix used for our analysis, the upper bound for the overlap between any symmetric matrix and the structural connectivity matrix was $(1 - N_a/2N_c) \approx 0.88$, where $N_a = 120$ is the number of asymmetric connections in the structural connectivity matrix and $N_c = 505$ is the total number of connections.

On top of the symmetry, there are further limitations on the functional connectivity matrix. In particular, there are some limitations on the correlation matrix stemming from the positive semi-definiteness of any covariance matrix. The simplest example is a sort of ‘weak’ transitivity - for any three random variables X, Y, Z a strong positive correlation between two pairs of them implies a positive correlation within the third pair. More precisely, $\rho_{XY}^2 + \rho_{YZ}^2 > 1$ implies $\rho_{XZ} > 0$, i.e. positivity of the third correlation coefficient (for a proof of a general form of this inequality see [105]). In practical terms this relates to the already reported fact that functional connectivity is sensitive to indirect connections. The level of this sensitivity has not been studied in detail and is likely to depend on the form of the underlying dynamical system and specific functional connectivity measure used.

Problems such as those outlined suggest that it might not be most suitable to compare structural connectivity with functional connectivity directly such as in

[82]. More adequate appears to be the approach adopted more recently in [84], where the observed human fMRI functional connectivity was compared with the functional connectivity predicted by a computational model that used human structural connectivity matrix for the inter-modular coupling. Clearly, such a model requires at least a basic model of haemodynamic response to model the process of conversion of neural activity to BOLD signal temporal variations. In the context of comparison between structural and functional connectivity it is also important to consider the distinction between *functional connectivity* and *effective connectivity* as introduced in [56] and discussed in section 3.3.2. A careful consideration of the distinction provides support for the notion that effective connectivity should be more closely than functional connectivity linked to the underlying structural connectivity. Apart from potential use of systematic perturbations of the system, causal relations implied in effective connectivity can be in principle studied in time-series due to temporal precedence of causes before effects. Particular methods include causality measures such as Granger causality [70], conditional mutual information (widely known under the name transfer entropy [157]) and fitting parameters of a dynamical model of a more constrained form that uses more prior information about the system. A prominent example of the latter approach to effective connectivity extraction is DCM [60]. The notion that effective connectivity measures might be linked more closely to the structural connectivity is supported by an observation from our model. The SC/FC overlap yielded generally higher values when TE was used instead of simple linear correlations - compare the Figures 9.5 and 9.3. While this effect might be partly due to the nonlinearity of TE, the crucial advantage of TE might also stem from the fact that it yields generally asymmetrical connectivity matrices.

When considering various functional or effective connectivity measures available, one should bear in mind that they provide different information about the underlying system, and their suitability may depend on the context. This context can be seen to have two main aspects. The first important issue is the nature of the data available. Neuroimaging data are collected by very different technologies and with wide range of settings and preprocessing approaches. At the very least, this yields time series with variable temporal and spatial resolution, intrinsic smoothness, distribution and dimensionality. Thus, connectivity measures useful for some data may not be so suitable for other datasets. While maybe less overt, this is also true about the structural connectivity - an evident example is the limitation of typical structural connectivity matrices yielded by

diffusion weighted imaging, which are not able to distinguish the direction of the fibre connection - in contrast to some invasive tracing approaches.

The second point of view is the question of the aim of the connectivity study. It is important to ask the question of what exactly do we actually look for when we study functional connectivity. To illustrate this issue we suggest the following simplified picture: two extreme approaches to functional connectivity could be roughly described as ‘utilitarian’ versus ‘theoretical’. The utilitarian approach follows the promise functional connectivity holds for applications. As mentioned earlier, changes in functional connectivity have been reported in many patient groups, suggesting potential for diagnostic use of functional connectivity. Similarly, functional connectivity may prove very useful for presurgical mapping of areas of so-called ‘eloquent’ cortex. Examples of the research promising such applications can be found e.g. in a recent review [175]. From a clinician’s point of view, it may not be as important to know what the functional connectivity measure represents, how it is related to structural connectivity and the nature of the underlying system and the data acquisition method – as long as the measure proves useful for the goals of interest such as diagnosis. On the other hand, the theoretical approach sees the challenge of functional connectivity study in the question ‘What can be learned about the underlying system (brain) from the spatiotemporal structure of its measured activity’. To study this question, models of the underlying system inevitably have to be constructed. Despite the necessary simplification and abstraction of variable extent, at least in principle the variables in the model should have identifiable counterparts in biophysical quantities describing the real brain.

While these outlined approaches may seem to aim in different directions, in practice they are intertwined. The basic research into brain dynamics is strongly motivated, supported and funded due to the aim of clinical applications. On the other hand, the search for the most useful measures of functional connectivity calls for deeper understanding of the underlying systems. Naturally, where the theoretical understanding is not yet available, pragmatic engineering choices have to be made.

It is valuable to ponder that the already widely used functional and effective connectivity measures already represent some strong theoretical assumptions, which can be related to very specific underlying models. Only understanding the underlying models allows proper interpretation of the coefficients or matrices that form output of these methods.

DCM is in this way the least cryptic, even containing the word ‘modelling’ in the name of the method. The basic bilinear dynamical system model is also coupled with a detailed model of the haemodynamic response, giving the model more biophysical realism. Note that DCM has not been widely applied on resting state data, also because in its current formulation it requires a definition of a stimulation time series. In other words, the method is not intended for direct application to intrinsic dynamics and would require some amendments.

An excellent example of a ‘hidden’ model is Granger causality. The most common implementation actually involves fitting a linear autoregressive model to the data and testing the hypothesis that some of the lagged linear regression coefficients are zero. This model implicitly assumes that the data are autocorrelated, and have a structure of stationary linear, but not necessarily symmetric interactions, which happen on a timescale comparable to that of the sampling frequency. Direct application of the method further assumes that the observable quantities do not suffer from different lags in different areas due to spatially variable haemodynamic response - a hidden model assumption that forms one of the main complications for use of Granger causality on fMRI BOLD data. A recent heated debate of the consequences of the different set of model assumptions underlying DCM and GC highlighted the importance of an understanding of the background of the connectivity models [57, 153].

As a third and last example, we note that even the most common linear correlation as an functional connectivity measure effectively reflects an underlying assumption of bivariate normality and lack of temporal autocorrelation of the time series. While the correlation might still bear valuable information about the dependence structure even under deviation from these assumptions, strong violation of these assumptions would render linear correlation suboptimal or even misleading as a measure of connectivity.

9.4 Modelling LFF

A second central aspect of spontaneous brain activity based on the neuroimaging observations is the low-frequency fluctuation of signals. While this LFF is directly observed in low-temporal resolution (~ 1 Hz) fMRI signal, they have been proposed to correspond to fluctuation of band-power of the underlying fast oscillations in electrophysiological signals [110, 118, 127]. As discussed earlier, the computational models of spontaneous brain activity presented in section 8.5

have not satisfactorily captured this behaviour.

In the model presented in this chapter, the time courses of activity of the cortical populations correspond to sustained fast oscillations of apparently stable amplitude and frequency (Figure 9.2). Nevertheless, the model activity is not completely periodic. In the following the temporal dynamics of the model are explored in a little more detail and potential emergence of low-frequency modulation of the fast oscillations is investigated.

While the individual areas slightly differ in their typical frequency due to varying level of incoming connections, for a given parameter setting the frequencies are in a relatively narrow band. For the following analysis we choose the parameter setting $P = -2, Q = -6, \gamma = 0.01$. Our particular parameter choice leads to oscillations with frequency around 19Hz (18.8 – 19.1Hz), when the time unit is set to 10ms. With these parameter settings, 1000 seconds of the system behaviour were simulated with a sampling step $\Delta t = 1\text{ms}$. In the first analysis, we focus on the fluctuation of amplitude of a single module. To this end, the single module time course was first band-pass filtered with a third order Butterworth filter with a pass-band 18.6-19.2Hz. Subsequently, the Hilbert transform was used to extract the amplitude of the band-pass filtered signal. Finally, the amplitude spectra of the resulting amplitude ‘envelope’ was computed. This amplitude spectra shows clear prevalence of low frequencies ($< 1\text{Hz}$) with highest amplitude for frequencies $< 0.1\text{Hz}$ (see Figure 9.6). While such spectral distribution corresponds to the observations of band-power ‘envelope’ spectra [111], the actual amplitude of the oscillation shows only very weak fluctuations, as evident from Figure 9.7. Such stability is in contrast with typical electrophysiological recordings, which show waxing and waning oscillations.

Importantly, a more biologically realistic picture is obtained when the average activation of the excitatory populations $\bar{E}(t) = 1/N \sum_{i=1}^N E_i(t)$ is considered. The spectrum of the envelope of the band-passed filter mean signal $\bar{E}(t)$ is shown in Figure 9.8. The dominance of the ultra-slow frequencies (0.01 – 0.1Hz) is even clearer than in the case of a single module. Also, in the case of mean signal, the amplitude of the oscillations shows remarkable fluctuations (Figure 9.9). Such temporal dynamics strongly resemble the dynamics of EEG amplitude including the long-temporal correlations *and* the strong variability in the amplitude [111].

This observation suggests that a potential mechanism for the emergence of LFF in the EEG oscillation amplitude or the downstream fMRI signal might be related to averaging over a big set of neural oscillators. This needs to be consid-

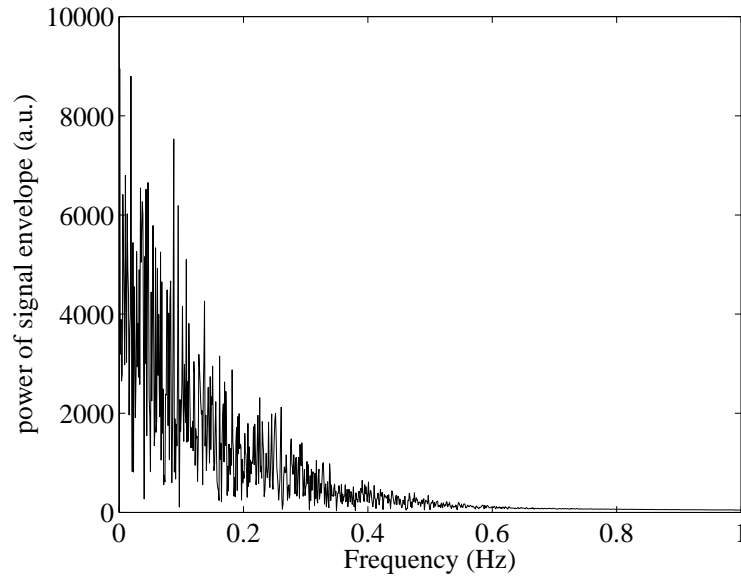


Figure 9.6: The amplitude spectra of the E variable of a single module of the Wilson-Cowan network. The parameters were set to $P = -2, Q = -6, \gamma = 0.01$.

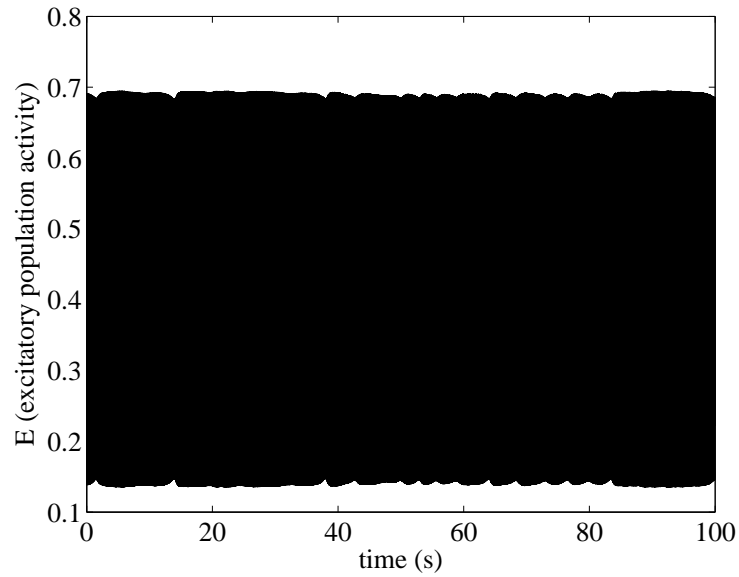


Figure 9.7: The time-course of the E variable of a single module of the Wilson-Cowan network. The parameters were set to $P = -2, Q = -6, \gamma = 0.01$.

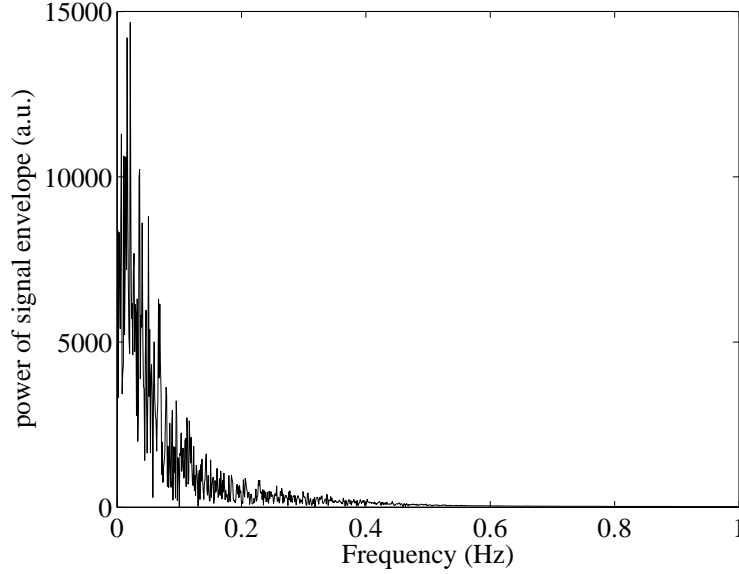


Figure 9.8: The amplitude spectra of the \bar{E} variable - the excitatory module activity averaged over all Wilson-Cowan nodes of the network. The parameters were set to $P = -2, Q = -6, \gamma = 0.01$.

ered as a rather abstract principle, as the model shown does not implement a sufficient level of biological detail. The most obvious problem disqualifying the straightforward interpretation is that in this *particular* model the Wilson-Cowan modules were originally considered as each representing a separate cortical area. Simple unweighted averaging of their signals therefore does not provide the most suitable model for the signal obtained from a specific electrophysiological sensor. Nevertheless, especially if the Wilson-Cowan modules were considered as models of cortical columns (which is also closer to the original Wilson-Cowan model assumptions [178]) the interpretation of suitably weighted averaged signal as a proxy for an electrophysiological recording would gain in relevance. While the original CoCoMac connectivity matrix loses its anatomical support in this context, we conjecture that the qualitative results will not be crucially dependent on a particular connectivity matrix used and (and particular averaging weights) as long as the main properties of the model are reproduced.

9.5 Summary

In the recent years, the large-scale spontaneous brain activity research matured to a stage where first explicit mathematical models of the structure and dynam-

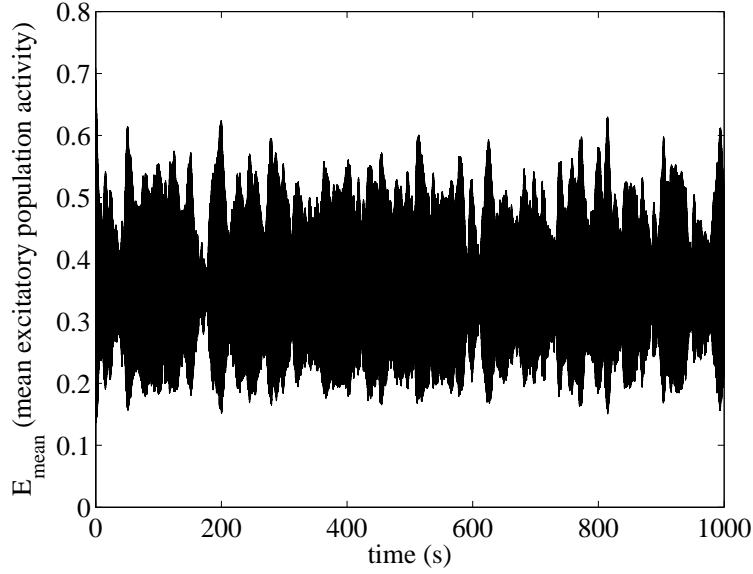


Figure 9.9: The time-course of the \bar{E} variable - the excitatory module activity averaged over all Wilson-Cowan nodes of the network. The parameters were set to $P = -2, Q = -6, \gamma = 0.01$.

ics of the system are being formulated. In this chapter we have carried out an analysis of such a basic inter-regional model of spontaneous brain activity from several aspects. The main purpose was to document some of the challenges in resting state modelling. Firstly, we have discussed the relation between the underlying structural and the generated functional connectivity. Importantly we have shown that this relation depends strongly on the parameters of the underlying model, even when the model and structural connectivity matrix is kept constant. This basic observation suggests possible problems with computational models when the parameters lack direct biophysical interpretation and therefore their choice remains largely arbitrary. This may also explain in part the difference in the structure/function relationship observed even in studies carried out by the same group [82, 84].

In general, any computational models of spontaneous brain activity come across interpretation challenges. We propose that as close linking to measurable quantities as possible is crucial to overcome these problems.

This is also important in the second central topic we have touched in this chapter – that is modelling of the LFF. Here we have shown that even a simple model can reproduce the slow modulations of the dominant fast cortical rhythm, corresponding to the LFF as discussed in the section 8.5. Contrary to the suggestions

in recent literature, this was possible even without explicit inclusion of noise and delays in the model. Importantly, averaging across populations was needed to show realistic level of temporal variability of the band-limited power of the signal. This motivates a question about the reasonability of the typical modelling choice of representing each cortical area by a single population or even a single neuron model. In the following chapter we focus in detail on an alternative model of a population of neurons that leads to the generation of nested fast and (ultra)slow oscillations.

Local model of LFF

As reviewed in Section 8.5 and further discussed in Chapter 9, current models of spontaneous brain activity have not yet fully addressed the question of the origin of LFF. Importantly, all of the proposed mechanisms rely on long-range inter-regional interactions, and some advocate the necessary role of transmission delays and noise. On the contrary, in this chapter we propose a local model of emergence of LFF within a local population of cortical neurons. The work presented in this chapter appeared in Physical Review Letters [80].

10.1 Introduction

The principle of the proposed model lies in postulating a local feedback loop regulating the activity level based on previous memory of the localised system. As an example of such a regulatory process we have implemented a simple phenomenological model of the action of endogenous cannabinoids on synaptic activity. Indeed, other known regulatory mechanisms could be also considered. Importantly we document that the local network activity can show slow to ultra-slow fluctuations that do not have to match the timescale of the memory mechanism in action. In principle, they can exhibit arbitrarily slow frequencies dependent on other parameters of the model.

Endogenous cannabinoids (CBs) represent a fundamentally new class of *retrograde* messengers [55], which are released postsynaptically and bind to presynaptic CB receptors. CB synthesis is stimulated when levels of calcium rise inside the neuron or when certain G-protein-coupled receptors are activated. One function of endogenous CBs is to regulate neurotransmitter release via activation of presynaptic CB receptors, allowing neurones to regulate, via feedback,

their upstream neuronal inputs [179]. This suppression of upstream presynaptic release of GABA or glutamate is termed *depolarisation-induced suppression of inhibition* (DISI) or *depolarisation-induced suppression of excitation* (DISE) respectively [137, 171]. Cannabinoid receptors are ubiquitous within the brain and CB1, the most abundant CB receptor, can be found in different areas such as the hippocampus, neocortex, amygdala, basal ganglia and hypothalamus [55]. They have already been implicated in the temporal coordination of cell assemblies and the modulation of certain brain rhythms [55, 152].

In the following section we endow a synaptically coupled network of Morris-Lecar neurons with a phenomenological form of retrograde second messenger signalling that can support DISE. The analysis in the subsequent section documents the emergence of nested fast and ultra-slow oscillations. We hypothesise that when linked to other modules in a larger network the latter would be reflected as an ultra-slow component of the macroscopic network dynamics and could therefore underlie those seen in spontaneous brain activity (SBA).

10.2 Model description

10.2.1 Synaptically coupled network of Morris-Lecar neurons

For the single neuron we have chosen the Morris-Lecar (ML) [130] neuron model. This is a classical two dimensional conductance based model, often used as an idealized fast-spiking pyramidal neuron, written in the form

$$\dot{v} = f(v, w) + I + s(t), \quad \dot{w} = g(v, w). \quad (10.1)$$

Here v plays the role of a voltage variable, w that of a gating variable, I is a fixed input and $s(t)$ represents a time varying synaptic input. The details of the functional forms for $f(v, w)$ and $g(v, w)$, which we take from [77] (with time measured in ms), have been reviewed previously in Section 8.2. The structure of the phase-plane and nullclines is recapitulated in Figure 10.1 for $s = 0$. In brief, the model can support either one or three fixed points, dependent on the choice of I . A crucial section of the bifurcation diagram is shown in the inset of Figure 10.1. This shows that the largest of the fixed points undergoes a sub-critical Hopf instability at $I \sim 0.0756$. Beyond this bifurcation there is a window of parameter space where a fixed point is bistable with a periodic orbit (see Section 8.2 including a more complete bifurcation diagram in Figure 8.5 for

additional details).

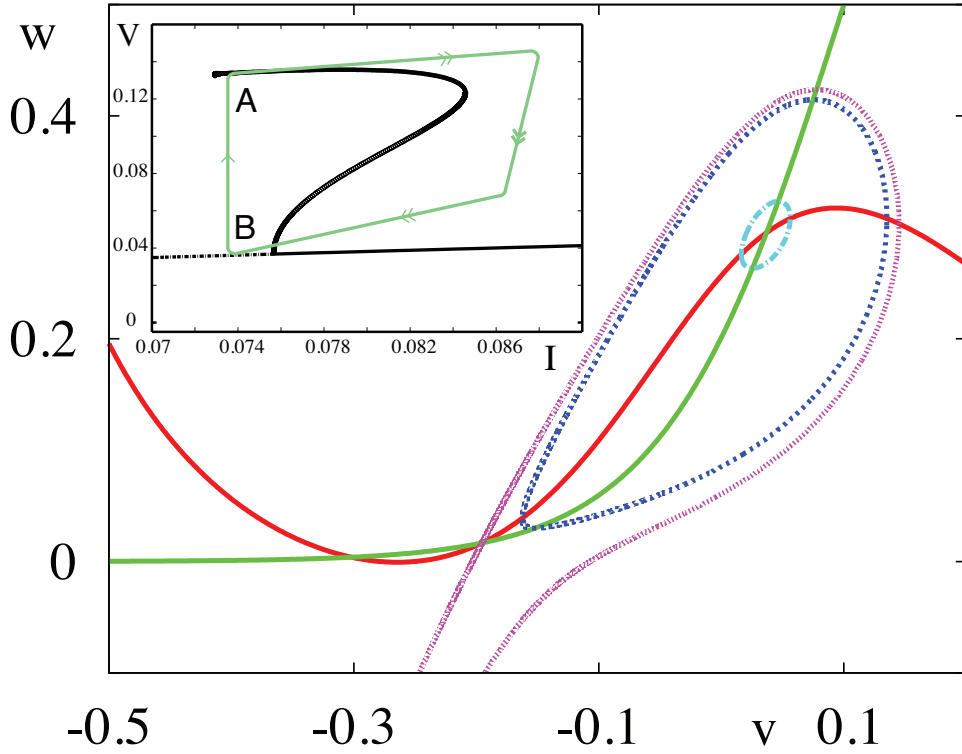


Figure 10.1: Phase-plane portrait for the Morris-Lecar model with constant external drive $I = 0.0761$. The voltage (gating) nullcline is in red (green). A large amplitude stable limit cycle (blue) coexists with a stable fixed point at $v \sim 0.04$. A small amplitude unstable orbit also exists (light blue) that can undergo a saddle-node of periodic orbits bifurcation with increasing I . The separatrix (pink, stable manifold of saddle at $v \sim -0.2$) delimits the basin of attraction for the stable fixed point at $v \sim -0.3$. The associated bifurcation diagram illustrating bistability of the large amplitude limit cycle and the fixed point at $v \sim 0.04$ is shown in the inset. Here unstable orbits emerge in a Hopf bifurcation.

Indexing each neuron in the network with $i = 1, \dots, N$ the synaptic drive to the i -th neuron is given by

$$s_i(t) = g_s(v_s - v(t)) \sum_{j=1}^N W_{ij} \sum_{m \in \mathbb{Z}} \eta(t - T_j^m), \quad (10.2)$$

where T_j^m is the m -th firing time of the j -th neuron, v_s the synaptic reversal potential and W_{ij} the connection strength between neurons i and j with a global conductance scaling g_s . The function $\eta(t)$ captures the shape of a conductance change in response to the arrival of an action potential. Here we choose an alpha function and write $\eta(t) = \alpha^2 t e^{-\alpha t} H(t)$, where H is a Heaviside step function.

The firing times are specified in terms of a threshold h according to

$$T_i^m = \inf\{t \mid v_i(t) > h, \dot{v}_i > 0, t > T_i^{m-1}\}. \quad (10.3)$$

We focus on the case of an excitatory globally coupled network with homogeneous connectivity and therefore set $W_{ij} = 1/N$ and $v_s = 2 > 0$ with respect to the resting state.

10.2.2 DISE mechanism

To implement a phenomenological model of DISE (or DISI if $v_s < 0$) we first introduce an effective spatio-temporal average level of depolarisation $v_e(t)$:

$$v_e(t) = \frac{1}{N} \sum_{j=1}^N \int_0^\infty K(t-s) v_j(s) ds, \quad (10.4)$$

where K is a temporal kernel reflecting the cannabinoid dynamics, $K(t) = 0$ for $t < 0$. Since CBs are free to diffuse through neural tissue they allow for *volume signalling*, which we will consider as a form of global feedback. The endocannabinoid level is directly linked to the effective depolarisation, which we shall take to be the average tissue level as defined by (10.4). Here we choose $K(t) = \lambda e^{-\lambda t} H(t)$, where λ^{-1} is an indirect measure of the long time-scale for cannabinoid dynamics, which is on the order of tens of seconds to minutes [171]. As a minimal model of DISE we will imagine that if the global CB level is sufficiently high then all excitatory synapses are blocked. In this case the network becomes uncoupled in the sense that excitatory synaptic currents drop to zero. By noting that this is equivalent to the suppression of firing events in (10.2), we may implement this model of DISE by letting the firing threshold adjust in response to $v_e(t)$ according to

$$h = \begin{cases} v^{\text{th}} & v_e \leq v_e^{\text{th}} \\ \infty & v_e > v_e^{\text{th}} \end{cases}. \quad (10.5)$$

The threshold v_e^{th} controls when the level of CB is sufficient to trigger DISE. In essence the model (10.5) means that synaptic interaction is curtailed if the effective level of depolarisation becomes too large.

10.3 Results

10.3.1 Model properties

To probe the effects of DISE on network dynamics we first focus on the most symmetric oscillatory states expected to exist in a globally coupled system – namely the fully synchronous and asynchronous ‘splay’ (evenly distributed) solution. These are guaranteed by symmetry arguments [5].

Synchronous solutions

In the synchronous state all neurons have identical T -periodic trajectories with firing times given by $T_i^m = mT$ for all i . In this case the synaptic drive to every neuron takes the identical form $s(t) = g_s(v_s - v(t))P(t)$, where $P(t) = \sum_{m \in \mathbb{Z}} \eta(t - mT)$ can be shown to equal to

$$P(t) = \frac{\alpha^2 e^{-\alpha t}}{1 - e^{-\alpha T}} \left[t + \frac{T e^{-\alpha T}}{1 - e^{-\alpha T}} \right], \quad t \in [0, T), \quad (10.6)$$

with $P(t)$ periodically extended outside $[0, T)$. Equation (10.1) may then be solved as a periodic boundary value problem (PBVP) for the periodic orbit $(v(t), w(t)) = (v(t+T), w(t+T))$ with $v(0) = v^{\text{th}}$. This describes the synchronous orbit given that the corresponding mean depolarisation does not trigger the DISE mechanism.

We solve this PBVP numerically, using XPPAUT[46] – see the source code in Appendix A. As mentioned earlier, existence of the synchronous solution for weak coupling is guaranteed by symmetry arguments. The period of these solutions as a function of the coupling g_s is shown in the Figure 10.2.

Whilst these synchronous solutions must exist for small enough g_s , using a weakly coupled oscillator description with standard techniques described in section 8.4 and also reviewed in [85], we further establish that such solutions are unstable. As mentioned in section 8.4, the stability of a phase-locked solution of the system of oscillators depends on the eigenspectra of the Jacobian \hat{H}_{ij} (8.40) of the linearized perturbation equation (8.38). In particular, while one eigenvalue is always zero, the solution is stable if all the remaining eigenvalues have negative real parts.

For the synchronous solution we have $\forall i, j \in \{1, \dots, N\} : \phi_i - \phi_j = 0$. Substi-

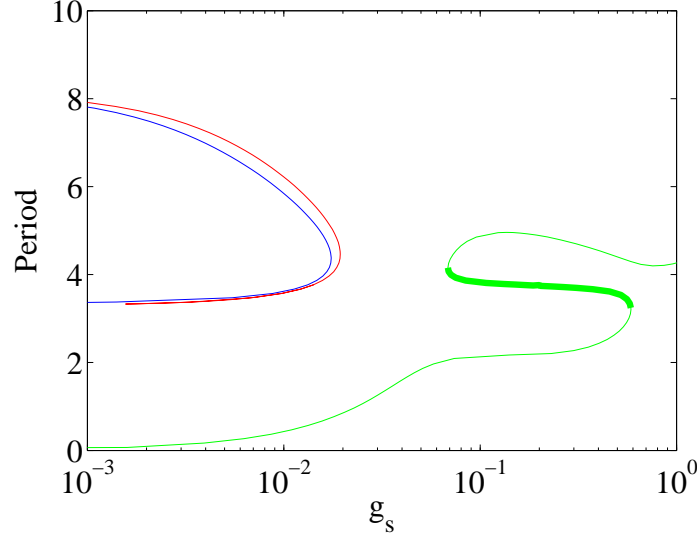


Figure 10.2: The period of the synchronous and splay solutions as a function of the coupling strength g_s . The unstable synchronous oscillatory solution for weak coupling is shown in blue, the synchronous solution branch for strong coupling is shown in green (numerically stable section in thick line). The unstable splay solution is shown in red.

tuting into (8.40) we obtain

$$\hat{H}_{ij} = \frac{1}{N} H'(0) - \delta_{ij} H'(0) = H'(0) \begin{pmatrix} \frac{1-N}{N} & \frac{1}{N} & \cdots & \frac{1}{N} \\ \frac{1}{N} & \frac{1-N}{N} & \cdots & \frac{1}{N} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{1}{N} & \frac{1}{N} & \cdots & \frac{1-N}{N} \end{pmatrix}. \quad (10.7)$$

It is easy to show that the latter matrix has the eigenspectra

$$\sigma = (0, -1, -1, \dots, -1)$$

with an $(N-1)$ -fold eigenvalue -1 . Therefore, the synchronous solution is stable if the $H'(0) > 0$. To check this condition, we generate the phase response curve $R(\theta)$ of the isolated Morris-Lecar neuron using XPPAUT [46] and convolve it after time-reversal according to (8.36) with the synaptic input (10.6). We obtain the phase interaction function H plotted in Figure 10.3. Clearly, the derivative of the function at zero is $H'(0) < 0$. Therefore, the synchronous solutions are unstable.

Note that another branch of synchronous solutions exists for a wide window of

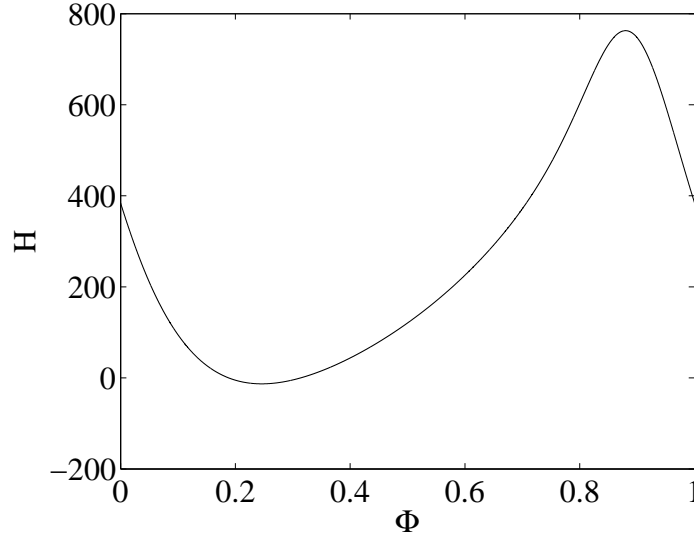


Figure 10.3: The phase interaction function of the synaptically coupled Morris-Lecar neurons.

g_s . This branch of solutions, distinct from the one corresponding to perturbation of the single neuron oscillatory solutions by a weak synaptic coupling, undergoes two turns, and putatively stable solutions are observed in the numerical simulations (implemented in Python using the Brian module [69] and in MATLAB) between these two turns (shown in thick green line in Figure 10.2). The Python source code is included in Appendix A.

Splay solutions

For an asynchronous splay state the firing times are given by $T_j^m = mT + jT/N$. In the limit $N \rightarrow \infty$ network averages may be replaced by time averages due to:

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{j=1}^N F(jT/N) = \frac{1}{T} \int_0^T F(t) dt, \quad (10.8)$$

for any T -periodic function $F(t) = F(t + T)$. Hence a splay state in which all neurons fire is given by $v_i(t) = v(t + iT/N)$, where $v(t)$ is a T -periodic solution of (10.1) with $s(t) = g_s(v_s - v(t))P_0$ and $P_0 = \int_0^T P(t) dt / T = 1/T$.

Solving the PBVP (source code in Appendix A), we find that for small g_s the splay state has a similar period to that of the synchronous solution. The period of these splay solutions is shown also in the Figure 10.2 (red line). Note that for the splay state $v_e(t)$ takes on the constant value $v_0 = \int_0^T v(t) dt / T$. This is lower

than the DISE threshold and therefore the DISE mechanism is not triggered. A weak-coupling analysis shows that the splay solution is also unstable – see the derivation below.

The phase differences between the oscillators $\phi_i - \phi_j$ span the values

$$0, \frac{1}{N}, \frac{2}{N}, \dots, \frac{N-1}{N},$$

leading to Jacobian

$$\hat{H}_{ij} = \frac{1}{N} \begin{pmatrix} -\sum_{i=1}^{N-1} H'(\frac{i}{N}) & H'(\frac{1}{N}) & H'(\frac{2}{N}) & \dots & H'(\frac{N-1}{N}) \\ H'(-\frac{1}{N}) & -\sum_{i=1}^{N-1} H'(\frac{i}{N}) & H'(\frac{1}{N}) & \dots & H'(\frac{N-2}{N}) \\ H'(-\frac{2}{N}) & H'(-\frac{1}{N}) & -\sum_{i=1}^{N-1} H'(\frac{i}{N}) & \dots & H'(\frac{N-3}{N}) \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ H'(-\frac{N-1}{N}) & H'(\frac{N-2}{N}) & H'(\frac{N-3}{N}) & \dots & -\sum_{i=1}^{N-1} H'(\frac{i}{N}) \end{pmatrix}. \quad (10.9)$$

This matrix is a Toeplitz matrix of a special form called a *circulant matrix*. A nice analytical result shows that the eigenvalues of a circulant matrix are given by Fast Fourier Transform of the first column of the matrix [37]. Indeed, the eigenvalues are

$$\lambda_n = \frac{1}{N} \sum_{j=1}^N H' \left(\frac{j}{N} \right) \left(e^{2\pi i \frac{nj}{N}} - 1 \right) = \quad (10.10)$$

$$= -\sum_{j=1}^{N-1} H' \left(\frac{j}{N} \right) \cdot 1 + \sum_{j=1}^N H' \left(\frac{j}{N} \right) \left(e^{2\pi i \frac{nj}{N}} \right). \quad (10.11)$$

To sidestep the numerical computation of derivative of the phase interaction function, we can use the fact that the Fourier coefficients of the derivative are linked to the Fourier coefficients of the original function, giving in the limit of large N

$$\lambda_n = -2\pi i n H_n, \quad (10.12)$$

where H_n are the Fourier coefficients of $H(\theta) = \sum_n H_n e^{2\pi i n \theta}$. These Fourier coefficients H_n can be further conveniently computed from the Fourier coefficients of the synaptic input $P(\theta)$ and phase response function $R(\theta)$ as shown below.

Consider the Fourier series

$$R(\theta) = \sum_n R_n e^{2\pi i n \theta}, \quad P(\theta) = \sum_n P_n e^{2\pi i n \theta}. \quad (10.13)$$

It can be easily shown from the definition (8.36), using $\int_0^1 d\theta e^{2\pi i (n+m)\theta} = \delta_{n,-m}$, that for the coefficients H_n of the series

$$H(\theta) = \sum_n H_n e^{2\pi i n \theta} \quad (10.14)$$

it holds that $H_n = P_n R_{-n}$.

Since we use the synaptic response in a relatively simple analytical form of an alpha function, we can obtain the P_n coefficients in an analytical form as $P_n = \alpha^2 \Delta / (\alpha \Delta + 2\pi i n)$, where Δ is the period of the oscillation. Therefore, the only numerical operation necessary is computation of the Fourier coefficients of the Phase Response Curve $R(\theta)$ of the Morris-Lecar neuron.

Carrying out the outlined computations, we finally obtain the eigenvalues given by (10.12). The real parts of the first 50 eigenvalues are plotted in Figure 10.4. Clearly, after the initial zero eigenvalue there are several eigenvalues with positive real part, confirming the instability of the splay solution.

The computations necessary to check the stability of both the splay and synchronous solution were carried out in MATLAB. See Appendix A for the relevant source code.

Clustered solutions

Whilst we have shown by theoretical arguments the instability of both the synchronous and splay solution, direct numerical simulations of the network suggest that another specific stable oscillatory solution exists even for the weak coupling. This has a hybrid form where the network splits into several clusters of fully synchronised neurons. These clusters then form a splay with evenly distributed phases.

Interestingly this type of solution typically further combines with a special type of clustered solution that can also occur for a wide range of g_s . This type of

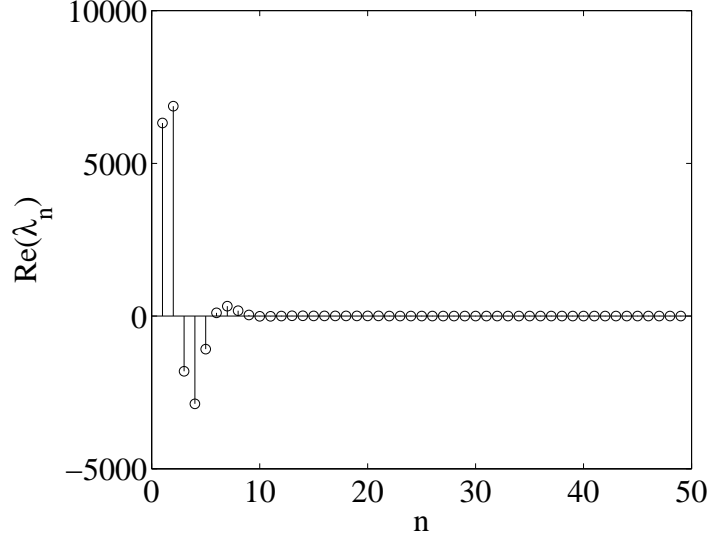


Figure 10.4: The real parts of the first 50 eigenvalues of the Jacobian \hat{H}_{ij} of the linearised perturbation system of the splay phase-locked solution of the synaptically coupled Morris-Lecar neuron network. The existence of eigenvalues with positive real parts proves the instability of the splay solution.

solution can be predicted purely from the theory as described below.

Consider two clusters of neurons. One is in a (clustered) splay state, populating the orbit corresponding to the stable limit cycle of a single neuron. The other cluster consists of neurons sitting at rest at the central fixed point, which is stable for a sufficient level of synaptic input from the oscillating cluster of neurons. This can be described using the differential-algebraic system

$$\begin{aligned} \dot{v} &= f(v, w) + I + r \frac{g_s}{T} (v_s - v), & \dot{w} &= g(v, w), \\ 0 &= f(\bar{v}, \bar{w}) + I + r \frac{g_s}{T} (v_s - \bar{v}), & 0 &= g(\bar{v}, \bar{w}), \end{aligned} \quad (10.15)$$

where r is the fraction of firing neurons and $(v(t+iT/M), w(t+iT/M))$ with $M = Nr$, and (\bar{v}, \bar{w}) describe neurons in the splay and resting cluster respectively. In this case $v_e = rv_0 + (1-r)\bar{v}$.

For $v_e < v_e^{\text{th}}$ the parameter region of existence for such a solution is illustrated in the inset of Figure 10.5, where a pair of splay states (with $r \neq 1$) only coexists with a rest state for $rg_s \in [L, H]$. Here the splay state is annihilated in a saddle-node bifurcation at $rg_s = H$, while below $rg_s = L$ the central fixed point becomes unstable (assuming the oscillating cluster sitting at the upper branch of the limit cycle solution). For fixed rg_s , as g_s is increased, v_e grows until it reaches v_e^{th}

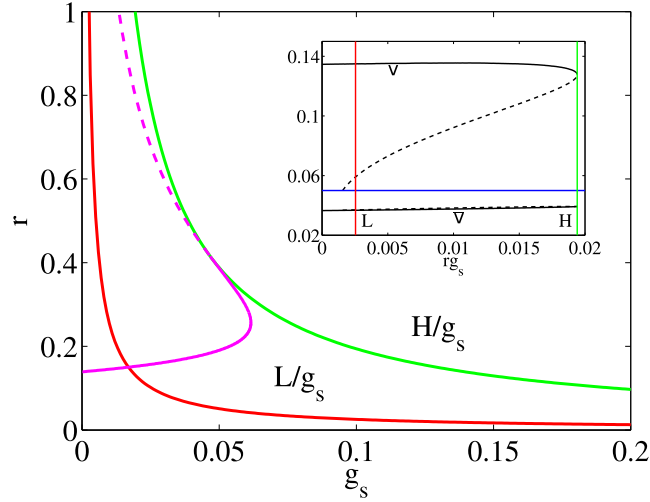


Figure 10.5: Fraction of firing neurons r as a function of the synaptic coupling strength g_s . The inset shows the bifurcation diagram (in maximum amplitude) for a splay state with $v_e < v_e^{\text{th}}$ as a function of the product rg_s . The rest state \bar{v} is less than the threshold $v^{\text{th}} = 0.05$ (blue line). Two branches of splay solution corresponding to stable (solid line) and unstable (dashed line) limit cycle of a single neuron model coalesce in a saddle-node of periodics at $rg_s = H$. L defines the loss of stability of the rest state under synaptic activity corresponding to the upper limit cycle branch. In the main figure the unlabelled magenta curved line shows the values of r for which $v_e = v_e^{\text{th}} = 0.02$. The green and red lines define the band of feasible values of r as a function of g_s corresponding to the splay cluster living on the larger orbit.

and activates the DISE mechanism. The border in the (r, g_s) parameter plane where $v_e = v_e^{\text{th}}$ for a cluster state is shown in Figure 10.5 (magenta line), and we see that it defines a critical curve marking the onset of DISE which we can write in the form $g_s = g_c(r)$. In the absence of DISE, cluster states with limit cycle corresponding to the upper branch of the limit cycle solution would exist for a greater area of parameter space defined by the right-infinite strip between the lines L/g_s and H/g_s .

For $g_s < g_c(r)$ direct numerical simulations do indeed show cluster states with properties in excellent agreement with the solution of (10.15) (with $v(0) = v^{\text{th}}$) up to small fluctuations. An example is shown in Figure 10.6.

For a given value of g_s the fraction of neurons r in the firing state is a function of initial data, as expected. Importantly, after transients, the mean depolarisation signal is flat (no oscillations) and the period of oscillation of a firing neuron is of the same order of magnitude as a single isolated neuron.

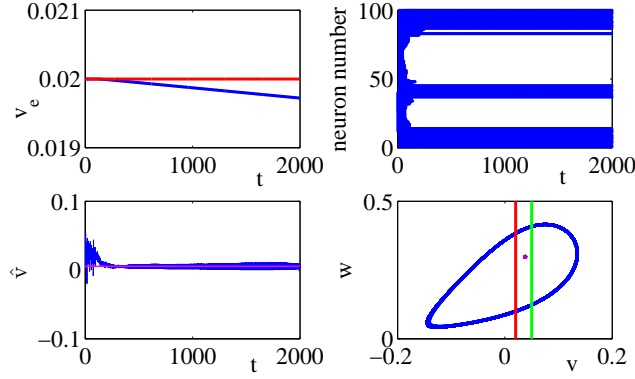


Figure 10.6: A cluster state for $N = 100$, $g = 0.03$, $v_e^{\text{th}} = 0.02$, $\lambda = 10^{-5}$. Top left: a plot of the average signal $v_e(t)$, showing that after transients the emergent state lies below the threshold to activate DISE (red line). Top right: A raster plot of spike times, illustrating the drop-out of some neurons and the emergence of a splay state with the fraction of firing neurons $r = 0.38$. Bottom left: The average network potential $\hat{v} = \sum_{i=1}^N v_i/N$ oscillates around the predicted value $rv_0 + (1-r)\bar{v}$ (magenta line) for $r = 0.38$. Bottom right: Phase plane dynamics for the network (dropping transients) showing that the network has split into two clusters (one with a common periodic orbit shown in blue with a period $T \sim 6$ and a rest state in purple). $v^{\text{th}} = 0.05$ (green line), $v_e^{\text{th}} = 0.02$ (red line).

10.3.2 Emergence of LFF

In the region where $g_s > g_c(r)$ and DISE precludes the existence of the above discussed cluster state we expect more exotic non-periodic network states to emerge. Notably, while stable synchronous oscillations are possible with increasing g_s , the average depolarisation for these rhythms is relatively high and also an increasing function of g_s . Hence there is also a critical value of g_s at which the DISE mechanism will also preclude the existence of this periodic *synchronous* state.

The mechanism for LFF emergence for this stronger coupling is as follows. A synchronous (or near synchronous) solution can lead to a strong level of average depolarisation for which $v_e(t) > v_e^{\text{th}}$. This activates the DISE mechanism, precluding further synaptic input and subsequently leading to a drop in network firing activity and hence a drop in $v_e(t)$. Once $v_e(t)$ drops sufficiently to cross v_e^{th} from above then excitatory synaptic currents can once again drive the network leading to an increase in $v_e(t)$ so that the process may repeat over. In this case the emergent time scale of the network rhythm is set by the duration of $v_e(t)$ above v_e^{th} . Even for a synchronous solution this will depend on initial data, so that network oscillations would not generically be periodic.

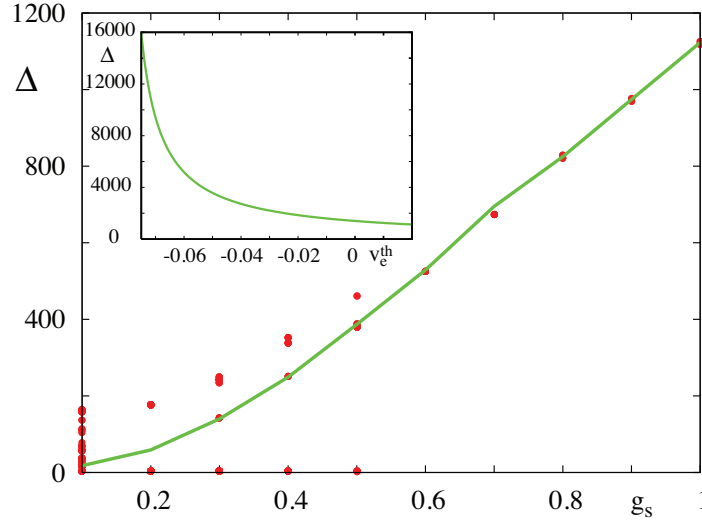


Figure 10.7: The predicted synchronous population ISI (in ms) as a function of g_s , for $w_0 = 0.121$ (green line), fits the ISIs seen in direct numerical simulations with $N = 100$ (red dots). Other parameters as in Figure 10.6. The inset shows the increase in ISI with decreasing v_e^{th} for $g_s = 1$.

To quantify the value of possible inter-spike intervals (ISIs) we focus on synchronous rhythms with $(v(0), w(0)) = (v^{\text{th}}, w_0)$ for some given w_0 and solve the BVP $v_e(0) = v_e^{\text{th}} = v_e(\Delta)$ with $s(t) = g_s(v_s - v(t))P(t)$. The relevant XPP source code is included in Appendix A.

The growth of the ISI, Δ , as a function of g_s is shown in Figure 10.7, together with results from direct simulations. The numerical spread of ISIs for low g_s can be ascribed to fast multi-spike bursts. With higher g_s a single spike response is more common and the period of the network state is accurately predicted by the theory.

Note that the spike times considered here are only those that contribute to synaptic currents, while the neurons do in fact spike on a fast time scale during the synaptically silent period. Hence, the network as a whole shows *nested* oscillations with a slow variation of synaptic currents superimposed on fast oscillations of the instantaneous average network voltage (see Figure 10.8 bottom left).

To understand how decreasing v_e^{th} can lead to a rapidly increasing Δ , as shown in the inset of Figure 10.7, it is useful to develop the correspondence of the evolution of the network (fixed parameters) with that of a single neuron with varying background drive I . Referring to the inset of Figure 10.1 the network can leave point A, corresponding to a synchronous firing state with average voltage v_2 , when $v_e(t)$ drops below v_e^{th} . The subsequent large increase in synaptic drive

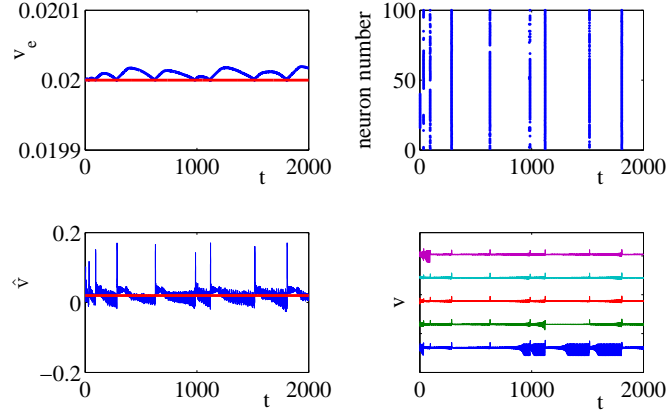


Figure 10.8: A similar plot to Figure 10.6 showing the emergence of slow synchronized firing patterns in the strong coupling regime with $g_s = 0.5$. Other parameters as in Figure 10.6. Bottom right shows voltage traces of 5 neurons (arbitrary offset for better display). Any variability due to heterogeneous initial conditions does not affect the emergence of ultra-slow near-synchronous oscillations.

causes a transition to the right of the saddle-node of periodics, where firing is not possible, and so synaptic currents fall which causes the transition to point B. This unstable fixed point, with voltage v_1 , generates orbits which spiral outward for a time $T_1 = T_1(g_s)$ generating a signal with $v_e(t) > v_e^{\text{th}}$ (so that synaptic currents are suppressed). These transition to full blown nonlinear oscillations, with average voltage v_2 and $\dot{v}_e(t) < 0$, and complete the path to point A so that the process may repeat over. Making the convenient (and obviously not accurate) assumption that $v_{1,2}$ are constant then the BVP may be solved by hand for $\lambda = 0$ to give $\Delta = T_1(g_s)(v_1 - v_2)/(v_e^{\text{th}} - v_2)$, explaining the dependence of Δ on v_e^{th} seen in Figure 10.7.

10.4 Discussion

Importantly, without any parameter fine-tuning, we see the emergence of very large ISIs for large values of g_s , which are largely independent of the network size. Moreover, in contrast to other network models of slow oscillations (< 1 Hz) [35] we do not require a mixture of excitation and inhibition, and as shown in the inset of Figure 10.7 with decreasing choices of v_e^{th} can easily achieve ISIs on the order of tens of seconds. Thus DISE in the strong coupling regime is a candidate mechanism for the generation of ultra-slow rhythms.

10.5 Summary

Low-frequency activity fluctuations constitute a characteristic feature of spontaneous brain activity. The recent pioneering models of large-scale spontaneous brain activity offer valuable contributions, but were so far not fully successful in suggesting or explaining mechanisms of the emergence of this feature. Moreover, these models often require the action of noise and signal transmission delays due to long-range cortico-cortical interactions.

In this chapter we have proposed a model for the emergence of low-frequency fluctuations in a localised neural population, independent of cortico-cortical delays or specific role of noise. The key mechanism involves a feedback loop regulating the activity level based on the previous states of the local system. We implement an example of such a regulatory process in a phenomenological model of the action of endogenous cannabinoids. These are retrograde messengers that suppress synaptic transmission under long-term increase in average depolarisation levels.

To understand the behaviour of the model, we have analysed its dynamics in detail, including the existence and stability of synchronous and asynchronous solutions in the limit of weak synaptic coupling. We have also observed existence of less symmetric but stable clustered solutions and described some of their properties. Most importantly, we document the emergence of solutions exhibiting nested fast and ultra-slow oscillations under relatively stronger synaptic coupling. This behaviour makes the action of such a regulatory loop a promising candidate for the mechanism of emergence of low-frequency fluctuations in spontaneous brain activity.

Discussion

In this thesis, I have described the state of the research of large-scale patterns of spontaneous brain activity. Special attention has been paid to the results stemming from the use of modern neuroimaging methods, in particular the functional Magnetic Resonance Imaging, and to computational neuroscience models that attempt to integrate the experimental evidence within a formal framework. The main current challenges of the field have been outlined, and I have presented several original research contributions directed towards addressing these challenges. Needless to say, this enterprise generated many new research questions, which are discussed in the following within the wider context of the field.

Confounding role of non-neuronal signals

The first original contribution of this thesis presented in chapter 4 was related to the potential confounding effects that increased head-movement can have on some indices of spontaneous brain activity. The main message of this chapter was the support for critical interpretation of the neuroimaging data. While the neuronal origin of *a part* of the BOLD fMRI signal has been firmly established over the last fifteen years, differences between conditions and groups may potentially arise from alterations in non-neuronal component of the signal.

Especially in a field without a strong underlying theoretical framework, where experimental results are likely to be interpreted *a posteriori* rather than against strict prior theoretical predictions, the danger of generating both false positive and false negative results cannot be overestimated.

It is pleasing to observe that more rigorous control of the equivalence of the potentially confounding variables such as motion is becoming increasingly adopted

in the most recent studies of effect of altered brain state of spontaneous brain activity features such as in [106].

Nongaussianity in fMRI functional connectivity

The question of appropriateness of the use of linear correlation for functional connectivity assessment in resting state fMRI is not only of abstract theoretical interest. To the opposite, the recent presentation of the message of chapter 5 at the Human Brain Mapping Conference (Barcelona 2010), confirming the very good fit of the gaussian approximation to the bivariate fMRI time series distribution, seemed to comfort the community of practical resting state fMRI researchers, worrying about potentially missed functional connections.

However, this result may be a bit confusing and disappointing to those more familiar with the electrophysiological and computational neuroscience research, where the nonlinear nature of the neuronal processes is highly acknowledged. This opens a range of relevant questions regarding the degree of potential generalisation of these results to other preprocessing methods and time-lagged dependencies. Explicit modelling might have something to say regarding the emergence of largely ‘linear’ or ‘gaussian’ dependencies on the macroscopic level of regional fMRI signals from the highly nonlinear nature of the small-scale neuronal dynamics.

Electrophysiological correlates of fMRI functional connectivity

The topic studied in detail in chapter 6 was the problem of existence of electrophysiological correlates of the fMRI functional connectivity. I have described an inter-subject relation between a slow EEG pattern and decreased functional connectivity in the default mode network, probably the most studied cortical resting state network. This can be seen as part of the general stream of research into resting state fMRI validation, focusing here directly on the functional connectivity construct rather than on the spontaneous fluctuation of the BOLD signal itself. On the other hand, this finding holds promise for potential relevance of fMRI FC for clinical (potentially diagnostic) application, since it documents that spontaneous functional connectivity has neuronally relevant inter-subject correlates.

These findings left several questions open. The most appealing is probably the question of the mechanism underlying the EEG-FC link. Only a speculative conceptual model was outlined in the respective chapter, but potentially in future work this question could be better addressed by development of dedicated models. Unfortunately, as discussed in the third part of this thesis, general models of large-scale spontaneous brain activity with sufficient predictive power are not yet readily available.

Challenges in spontaneous brain activity modelling

As described earlier, modelling large-scale spontaneous brain activity is a relatively new enterprise, still in the stage of initial attempts. This is reflected in the nature of the questions studied in the relevant chapters of this thesis – the focus is on the feasibility discussion and potential candidate mechanisms for some of the observed phenomena.

A careful reader may notice that a slightly different perspective on SBA modelling priorities is adopted from the one prevalent in the current literature, in focusing more on the emergence of the low-frequency fluctuations rather than reproducing the functional connectivity structure. The reason for this approach is two-fold. The first aspect is rather practical. The quality of predictions of functional connectivity may be currently seriously limited by the information available about the existence, directionality and relative strength of structural connections. Despite the improvements in the anatomical connectivity methods such as diffusion imaging based tractography, the sufficiency of the available data is still questionable.

The second reason is more principal. Functional connectivity represents nothing else than temporal synchronisation of activity of distant cortical modules. In this sense it can be considered secondary to the temporal dynamics, although this simplistic view has some limitations. Therefore, even a significant degree of recovery of synchronisation structure among cortical regions, when obtained from signals of grossly unrealistic temporal properties would raise questions about the wider interpretability of results of such a model.

The initial stage of the research in modelling of large-scale spontaneous brain activity phenomena is also reflected in the width of open questions revealed by the investigation in this thesis. From the richness of these open problems, I highlight two most closely related to potential future work on validation of

the proposed mechanism for low-frequency fluctuation emergence. First, further investigation is needed to reveal what is the role of regulatory feedbacks in the slow brain activity dynamics. The search should not be limited to the action of endocannabinoids that served as an example carrier of this mechanism - the role of other neuromodulators or metabolic fatigue should be investigated. On the other hand, attention should also be paid to the explanatory potential of the local synchronisation dynamics as a minimalistic model for low-frequency fluctuations. Here, a comparison with localised electrophysiological recordings might prove extremely beneficial.

Open questions in the broader context

The relatively wide scope of this thesis has inevitably led to some selectivity in the topics covered and limited detail devoted to particular areas. There are several areas I consider worth special attention of the community investigating the spontaneous brain activity, that could however be sufficiently represented in this thesis.

The first is the psychological and more widely functional relevance of the spontaneous brain activity. Indeed, there is a wealth of knowledge as well as methods available in the field of psychology that waits to be properly integrated with the newest neuroimaging results in spontaneous brain activity.

The second underrepresented area concerns the breadth of results of neurocomputational modelling mainly of electrophysiological signals that have been published and discussed rather in isolation from the newest observations in functional MRI. Moreover, general concepts such as dynamical synchronisation (see e.g. [19]), self-organised criticality [29] or many other may be closely relevant for the observed features of the resting state neuroimaging data. Integration of these concepts is likely to be fruitful, although critical assessment of all candidate mechanisms is necessary to assure predictive power and biological realism of the chosen models. The danger here is that with enough free parameters, models could fit well *some* of the observed features without bringing any more understanding of the underlying system.

Conclusion

I would like to conclude this thesis with a few notes on the most recent development in the field and its outlook. While it might go unnoticed at first, the study of spontaneous brain activity has undergone a major transformation over the last few years. To document this, several of the newest trends are outlined below.

Probably the most noticeable of these trends is the wide interest the topic has recently attracted in the general neuroscience and neuroimaging community. Moreover, this interest has further materialised into establishment of a huge publicly available database of 1000+ resting state fMRI datasets [16], a move that will without doubt further increase the activity and recognition of the field.

A related point is the ‘warming’ of the relations with the mainstream task-related neuroimaging studies community. In fact, integration of task-related activations and resting state networks is one of the most appealing movements in the neuroimaging field [165].

Importantly, the spontaneous brain activity research community has dealt with the critical voices from the more established disciplines by internalising the critical approach. Among other things, this has lead to several recent studies critically assessing the reliability of the resting state features [183, 160, 126]. This critical element is apparent also in the ongoing reformulation of the crucial concepts of the SBA study. For instance, the notion of resting state networks is becoming increasingly understood in the form of multi-level hierarchical networks. The level of detail observed then depends on solely on the method chosen – a view closer to the intuitive understanding of organisation of brain function. Another example is the recent upraise in the acknowledgement of the role of subcortical structures in the SBA, including cerebellum, basal ganglia and other structures. It is impossible to list here all the recent trends and advances – in the end, the field has been suggested to be undergoing a change resembling a paradigm shift [146].

Without further eluding to the discussion of the scientific paradigm in neuroimaging and neuroscience in general, I hope this thesis managed to build a rich picture of the spontaneous brain activity study as a flourishing research area. The field is definitely in a stage of rapid expansion and maturation as a descriptive discipline. Perhaps even more importantly, this process is accompanied by the increasing recognition of the important role of theoretical integration

of experimental findings, validated through predictions of formalised models.

Throughout my doctoral studies, it has been a great pleasure to be both a witness and an active participant in this adventurous scientific journey to the realm of spontaneous brain activity. Although the form of this report is at places necessarily quite technical and the level of detail might distract from the bigger picture, I hope that I have conveyed to the reader at least part of that pleasure of scientific discovery of one of the most intimate mysteries of human mind – the mystery of spontaneous activity.

APPENDIX A

Selected source code

XPP code for the PBVP for the ML neuron synaptically coupled network splay solution

```
# Boundary value problem for the Morris-Lecar synaptically coupled
  network splay solution

# parameters (can be used for numerical continuation of solutions)
p g=0.01
p I=0.075
p vth=0.05

# initial values
v(0)=0.05
w(0)=0.1222481
per(0)=6.22141

# constants/fixed model parameters
gca=1.0, gk=2.0, gl=0.5, vk=-0.7, vl=-0.5, f=1.15
v1=-0.01, v2=0.15, v3=0.1, v4=0.145, vs=2

# static equations
minf=0.5*(1.0+tanh((v-v1)/v2))
winf=0.5*(1.0+tanh((v-v3)/v4))
tauw=1.0/cosh((v-v3)/(2.0*v4))

# boundary conditions for the BVP
bndry v-vth
bndry v'-vth
bndry w-w'

# dynamical equations
per'=0
v'=(per)*(-gca*minf*(v-1.0)-gk*w*(v-vk)-gl*(v-vl)+I)+g*(vs-v)
w'=(per)*(f*(winf-w)/tauw)

# XPPAUT settings
@ xp=v, yp=w, xlo=-0.5, xhi=0.2, ylo=-0.1, yhi=0.5
@ bounds=100000, maxstore=100000, nmesh=200, total=1.0
@ method=runge-kutta, dt=0.001
done
```

XPP code for the PBVP for the ML neuron synaptically coupled network synchronous solution

```

# Boundary value problem for the Morris–Lecar synaptically coupled
  network synchronous solution

# parameters (can be used for numerical continuation of solutions)
p g=0.01
p I=0.075

# initial values
v(0)=0.05
w(0)=0.125
per(0)=5.9

# constants/fixed model parameters
gca=1.0, gk=2.0, gl=0.5, vk=-0.7, vl=-0.5, f=1.15
v1=-0.01, v2=0.15, v3=0.1, v4=0.145
vs=2, vth=0.05, alpha=1

# static equations
minf=0.5*(1.0+tanh((v-v1)/v2))
winf=0.5*(1.0+tanh((v-v3)/v4))
tauw=1.0/cosh((v-v3)/(2.0*v4))
PT=alpha*alpha*(exp(-alpha*theta*per)/(1-exp(-alpha*per)))*(theta*
  per+(per*exp(-alpha*per))/(1-exp(-alpha*per)))

# boundary conditions for the BVP
bndry v - vth
bndry v'- vth
bndry w-w'
bndry theta

# dynamical equations
per'=0
v'=(per)*(-gca*minf*(v-1.0)-gk*w*(v-vk)-gl*(v-vl)+I)+g*per*PT*(vs-v)
w'=(per)*(f*(winf-w)/tauw)
theta'=1

# XPPAUT settings
@ xp=v, yp=w, xlo=-0.5, xhi=0.2, ylo=-0.1, yhi=0.5
@ bounds=100000, maxstore=100000, nmesh=200, total=1.0
@ method=runge-kutta, dt=0.001
done

```

XPP code for the BVP for the ML neuron network solution with DISE

```

# Boundary value problem for the Morris–Lecar synchronous solution
  endowed with the DISE mechanism

# parameters (can be used for numerical continuation of solutions)
p w0=0.121
p g=0.1
p I=0.075
p veth=0.02

# initial values
v(0)=0.05
w(0)=0.121
per(0)=25

# constants/fixed model parameters

gca=1.0, gk=2.0, gl=0.5, vk=-0.7, vl=-0.5, f=1.15
v1=-0.01, v2=0.15, v3=0.1, v4=0.145
vs=2, vth=0.05, alpha=1

# static equations
minf=0.5*(1.0+tanh((v-v1)/v2))
winf=0.5*(1.0+tanh((v-v3)/v4))
tauw=1.0/cosh((v-v3)/(2.0*v4))
PT=alpha*alpha*(exp(-alpha*theta*per)/(1-exp(-alpha*per)))*(theta*
  per+(per*exp(-alpha*per))/(1-exp(-alpha*per)))

# boundary conditions for the BVP
bndry v - vth
bndry w-w0
bndry theta
bndry vv
bndry vv'-veth

# dynamical equations
per'=0
v'=(per)*(-gca*minf*(v-1.0)-gk*w*(v-vk)-gl*(v-vl)+I)+g*per*PT*(vs-v)
w'=(per)*(f*(winf-w)/tauw)
theta'=1
vv'=v

# XPPAUT settings
@ xp=v, yp=w, xlo=-0.5, xhi=0.2, ylo=-0.1, yhi=0.5
@ bounds=100000, maxstore=10000000, nmesh=200, total=1.0
@ method=runge-kutta, dt=0.001

done

```

MATLAB code for computation of the Phase Interaction Function of the synchronous ML network solution

```

function y=PIF(PRC, lengthEta, v, vs, alpha)

% Jaroslav Hlinka
% version from July 2010
%
% FUNCTION: computes Phase Interaction Function. Note that
%           derivation of PIF at 0
% determines stability of synchronous solution of network of weakly
%           coupled
%           oscillators
%
% INPUT:
% PRC – Phase Response Curve. Accepted in three column format as
%       exported
% from XPP: first column time, second V-component of PRC, third W-
%           component
% of PRC (disregarded)
% alpha – speed of synapse – typically 1
% lengthEta – length of synaptic impulse to consider (I guess now
%           high values more precise, but take longer to compute?)
% v – voltage traces in one period. Accepted in three column format
%     as exported
% from XPP: first column time, second V-component of PRC, third W-
%           component
% of PRC (disregarded)
% vs – shunt voltage
%
% NOTES:
% synapse of alpha function form implemented

%% computation
if nargin < 5
    alpha = 1;
end
[s1,s2] = size(PRC);
deltat = PRC(2,1)-PRC(1,1);
t = 0:deltat:lengthEta;
N = alpha^2.*t.*exp(-alpha*t);
R = PRC(:,2)';
if nargin < 4
    shunt = ones(s1,1);
else
    shunt = (vs-v(:,2))';
end
Rshunt = R.*shunt;
y = deltat*cconv(N,Rshunt(end:-1:1),s1);
%% plotting
figure;
plot(0:1/(s1-1):1,PIF,'k');
xlabel('Phi');
ylabel('H');
set(findall(gcf, 'Type', 'text', '-depth', inf), 'FontName', 'times', 'FontSize', 20)
set(findall(gcf, 'Type', 'axes', '-depth', inf), 'FontName', 'times', 'FontSize', 20)

```

MATLAB code for computation of the eigenvalues of the ML network splay solution Jacobian

```

function lambdas=eigenvalues_splay(PRC,v,vs,alpha,Ndistinct)

%% initialisation
if nargin < 5
    Ndistinct = ceil(size(PRC,1)/2);
end
if nargin < 4
    alpha = 1;
end

%% computation
newsamplepoints = PRC(1,1):(PRC(end,1)-PRC(1,1))/(2*Ndistinct-2):PRC
    (end,1);
PRC = interp1(PRC(:,1),PRC,newsamplepoints); % resample the PRC
v = interp1(v(:,1),v,newsamplepoints); % resample the voltage trace
shunt=vs-v(:,2);
Rns = fft(PRC(:,2).*shunt);
delta = PRC(end,1);
l = length(Rns);
if mod(l,2)==0
    Rns = Rns(1:(end-1));
end
RevertedRns = Rns([Ndistinct:-1:1,end:-1:(Ndistinct+1)]);
nvalues = (-Ndistinct+1: 1:Ndistinct-1);
Pnvalues = (alpha^2*delta)./((alpha*delta+2*pi*1i.*nvalues).^2);
Hnvalues = Pnvalues.*RevertedRns';
lambdas = -2*pi*1i*nvalues.*Hnvalues;

%% plotting
figure
stem(nvalues(nvalues>=0),real(lambdas(nvalues>=0)),'k');
xlabel('n');
ylabel('Re(\lambda_n)');
set(findall(gcf, 'Type', 'text', '-depth', inf), 'FontName', 'times', '
    FontSize', 20)
set(findall(gcf, 'Type', 'axes', '-depth', inf), 'FontName', 'times', '
    FontSize', 20)

```

Python code for the DISE-endowed Morris-Lecar synaptically coupled network simulation

```

#ipython -pylab DISE_global.py
#ipython -pylab -wthread DISE_global.py
#execfile('DISE_global.py') from python
'''

Morris-Lecar synaptically coupled network with Depolarisation
Induced Suppression of Excitation (DISE)
J HLinka July 2010
'''

import brian_no_units
from brian import *
import time
tic = time.time()

#parameters
J=0.075; gca=1.0; gk=2.0; gl=0.5; vk=-0.7; vl=-0.5; f=1.15; v1
    =-0.01; v2=0.15; v3=0.1; v4=0.145
vth=0.05; kappa=1.0; veth=0.02; vsyn=2.0; alpha=1.0
N=10
duration=20
#looping parameters
gmin=0.03
gstep=0
grange=1 # number of g samples

# initialisation
gsave=[]
Fractionsave=[]
Periodsave=[]
vefinalsave=[]
AverPeriodsave=[]

#main code in a loop over g values
for ig in range(grange) :
    clear()
    g=gmin+ig*gstep
    simulation_clock=Clock(.1)
    print 'g is %f' % g
    print 'N is %d' % N
    print 'dt is %f' % simulation_clock.dt
    print 'duration is %5.2f' % duration, ' time units'

    # The model
    eqs=Equations(''')
    dv/dt = -gca*minf*(v-1.0)-gk*w*(v-vk)-gl*(v-vl)+J+s*(vsyn-v)
            : 1
    dw/dt = f*(winf-w)/tauw : 1
    ds/dt = alpha*(-s+x) : 1
    dx/dt = alpha*(-x) :1
    dve/dt = v : 1
    minf = 0.5*(1.0+tanh((v-v1)/v2)) : 1
    winf = 0.5*(1.0+tanh((v-v3)/v4)) : 1
    tauw = 1.0/cosh((v-v3)/(2.0*v4)) : 1
    ''')

    def vdot(v,w,s) :
```



```

    minf = 0.5*(1.0+tanh((v-v1)/v2))
    winf = 0.5*(1.0+tanh((v-v3)/v4))
    tauw = 1.0/cosh((v-v3)/(2.0*v4))
    return -gca*minf*(v-1.0)-gk*w*(v-vk)-gl*(v-vl)+J+s*(
        vsyn-v)

def mynewthreshold (v, w, s, x, ve) :
    #we need a threshold function that only generates a spike
    upon crossing threshold for the first time
    #find all the neurons for which dv/dt > 0 and v below
    threshold on last step
    #return v>vth for this set only
    r=v[:] * 0 #zeros of size N
    vv=sum(ve)/(N*simulation_clock.t) #average voltage
    across neurons and time (since ve is cumulative
    voltage of each neuron)
    for i in range(len(v)) :
        vd = vdot(v[i],w[i],s[i])
        # check that voltage increasing, in last
        step below vth and average voltage now
        below veth
        if vd > 0 and v[i]-vd*simulation_clock.dt <
            vth and vv < veth:
            r[i]=v[i]>=vth
        else :
            r[i]=v[i]>1000
    return r

P=NeuronGroup(N,model=eqs,threshold=FunThreshold(
    mynewthreshold),clock=simulation_clock,order=2,freeze=
    True)
C=Connection(P,P,'x') #why 'x'? -
    variable to be kicked!
C.connect_full(P, P, weight=alpha*g/N)

# Initialization
    # non-initialised x has IC=0
# ICs random:
P.v=0.03+0.001*randn(len(P))
P.w=0.11+0.001*randn(len(P))
P.s=0.001*randn(len(P))
P.ve=veth+0.001*randn(len(P))

# Record the number of spikes and a few traces
tracev=StateMonitor(P,'v',record=True,clock=simulation_clock
)
tracew=StateMonitor(P,'w',record=True,clock=simulation_clock
)
traceve=StateMonitor(P,'ve',record=True,clock=
simulation_clock)
traces=StateMonitor(P,'s',record=True,clock=simulation_clock
)
M=SpikeMonitor(P)

run(duration)
tt = arange(simulation_clock.dt,simulation_clock.t,
simulation_clock.dt) # like range, but allows floats

figure(None, (10, 7.5), 100)
clf()

```

```

#calculate period - ISI of last neuron to fire
Period=0
Fraction=0
if len(M.spikes) < 2*N :
    print 'No spikes\n',
    subplot(224)
    plot()
else :
    index = M.spikes[-1][0]                # last spike
    neuron index
    Tend = M.spikes[-1][1]                # last spike
    time
    Tlast = 0

    for k in range(len(M.spikes)) : # finds the previous
        spike of the last spiking neuron
        if M.spikes[-1-k-1][0] == index :
            Tlast = M.spikes[-1-k-1][1]
            break
        else :
            Tlast=Tend
    Period = Tend - Tlast
    print 'Period is %f' % Period

#calculate fraction of superthreshold neurons
myspikes = [0]
myspikes = myspikes*N
startcoef=0.75
validstart=startcoef*duration
validspikes=[x for x in M.spikes if x[1]>validstart]

for k in range(len(validspikes)) :
    myspikes[validspikes[k][0]]+=1
sub = []
for k in range(len(myspikes)) :
    if myspikes[k] < 5 : # counts off all
        neurons with < 5 spikes
        sub.extend([k])

X=set(range(N))
Y=set(sub)
super = X-Y
Fraction = len(super)*1.0/N
print 'Fraction is %f' % Fraction
if len(validspikes)==0:
    AverPeriod=0
else:
    AverPeriod=(1-startcoef)*duration*N/len(
        validspikes)
print 'Average Period is %f' % AverPeriod
subplot(224)
st = []
sn = []
st = st + [float(a[1]) for a in M.spikes]
sn = sn + [a[0] for a in M.spikes]
plot(st,sn, '.')
index=[]
index=int(len(tracev[0][:])*0.75)

subplot(222)

```

```

plot(tracev[:,index:], tracew[:,index:], (vth, vth), (0, 0.7), '
      r-', (veth, veth), (0, 0.7), 'g-')
xlabel('v')
ylabel('w')

ve = mat(traceve[:,1:]).T*1.0/tt[0:size(traceve[:,1:],0)]
ve=sum(ve,0)/N

subplot(221)
plot(tt[0:len(array(ve.T))], array(ve.T))
plot(tt[0:len(array(ve.T))], [veth]*len(ve.T))
xlabel('t')
ylabel('ve')
ylim(ymax=vth)

print 've final value is %f' % ve.T[-1]

subplot(223)
if super == set([]) :
    ham=0
else :
    ham = min(super)
plot(tt[-1], tracev[ham][1:len(tt)], 'r-', tt[-1], [vth]*len(
    tt[-1]), 'g-')
xlabel('t')
ylabel_str = 'v[' + str(ham) + ']'
ylabel(ylabel_str)
ylim(ymax=0.3)

title_str = "g = " + str(g) + ", Period = " + str(Period) +
    ", Fraction = " + str(Fraction)
text(0.5, 0.95, title_str)

toc = time.time()
print toc-tic, ' seconds have elapsed'

gsave.append(g)
Fractionsave.append(Fraction)
Periodsave.append(Period)
vefinalsave.append(ve.T[-1])
AverPeriodsave.append(AverPeriod)

gfile=open('gsave.txt','w')
gfile.write(str(gsave))
gfile.close()
Fractionfile=open('Fractionsave.txt','w')
Fractionfile.write(str(Fractionsave))
Fractionfile.close()
Periodfile=open('Periodsave.txt','w')
Periodfile.write(str(Periodsave))
Periodfile.close()
vefinalfile=open('vefinalsave.txt','w')
vefinalfile.write(str(vefinalsave))
vefinalfile.close()
AverPeriodfile=open('AverPeriodsave.txt','w')
AverPeriodfile.write(str(AverPeriodsave))
AverPeriodfile.close()

```

Bibliography

- [1] S. Achard, R. Salvador, B. Whitcher, J. Suckling, and E. Bullmore. A resilient, low-frequency, small-world human brain functional network with highly connected association cortical hubs. *Journal of Neuroscience*, 26(1):63–72, 2006.
- [2] P. J. Allen, O. Josephs, and R. Turner. A Method for Removing Imaging Artifact from Continuous EEG Recorded during Functional MRI. *NeuroImage*, 12(2):230–239, 2000.
- [3] P. J. Allen, G. Polizzi, K. Krakow, D. R. Fish, and L. Lemieux. Identification of EEG Events in the MR Scanner: The Problem of Pulse Artifact and a Method for Its Subtraction. *NeuroImage*, 8(3):229–239, 1998.
- [4] J. Alstott, M. Breakspear, P. Hagmann, L. Cammoun, and O. Sporns. Modeling the Impact of Lesions in the Human Brain. *PLoS Computational Biology*, 5(6), 2009.
- [5] P. Ashwin and J. W. Swift. The dynamics of n-weakly coupled identical oscillators. *Journal of Nonlinear Science*, 2(1):69–108, 1992.
- [6] D. P. Auer. Spontaneous low-frequency blood oxygenation level-dependent fluctuations and functional connectivity analysis of the ‘resting’ brain. *Magnetic Resonance Imaging*, 26(7):1055–1064, 2008.
- [7] C. Babiloni, E. Cassetta, G. Binetti, M. Tombini, C. Del Percio, F. Ferreri, R. Ferri, G. Frisoni, B. Lanuzza, F. Nobili, L. Parisi, G. Rodriguez, L. Frigerio, M. Gurzi, A. Prestia, F. Vernieri, F. Eusebi, and P. M. Rossini. Resting EEG sources correlate with attentional span in mild cognitive impairment and Alzheimer’s disease. *European Journal of Neuroscience*, 25(12):3742–3757, 2007.
- [8] S. Balkan, N. Yaras, E. Mihci, B. Dora, A. Agar, and P. Yargicoglu. Effect of donepezil on EEG spectral analysis in Alzheimer’s disease. *Acta Neurologica Belgica*, 103(3):164–169, 2003.
- [9] R. M. Baron and D. A. Kenny. The Moderator-Mediator Variable Distinction in Social Psychological Research - Conceptual, Strategic, and Statistical Considerations. *Journal of Personality and Social Psychology*, 51(6):1173–1182, 1986.
- [10] C. F. Beckmann and S. A. Smith. Probabilistic independent component analysis for functional magnetic resonance imaging. *IEEE Transactions on Medical Imaging*, 23(2):137–152, 2004.

- [11] M. Bianciardi, M. Fukunaga, P. van Gelderen, S. G. Horovitz, J. A. de Zwart, K. Shmueli, and J. H. Duyn. Sources of functional magnetic resonance imaging signal fluctuations in the human brain at rest: a 7 T study. *Magnetic Resonance Imaging*, 27(8, Sp. Iss. SI):1019–1029, Oct 2009. International School on Magnetic Resonance and Brain Function, Erice, Italy, May 18-25, 2008.
- [12] P. Bickel and P. Bühlmann. What is a linear process? *Proceedings of the National Academy of Sciences of the United States of America*, 93(22):12128–12131, 1996.
- [13] J. R. Binder, J. A. Frost, T. A. Hammeke, P. S. F. Bellgowan, S. M. Rao, and R. W. Cox. Conceptual processing during the conscious resting state: A functional MRI study. *Journal of Cognitive Neuroscience*, 11(1):80–93, 1999.
- [14] R. M. Birn, J. B. Diamond, M. A. Smith, and P. A. Bandettini. Separating respiratory-variation-related fluctuations from neuronal-activity-related fluctuations in fMRI. *NeuroImage*, 31(4):1536, 2006.
- [15] B. Biswal, F. Z. Yetkin, V. M. Haughton, and J. S. Hyde. Functional connectivity in the motor cortex of resting human brain using echo-planar MRI. *Magnetic Resonance in Medicine*, 34(4):537–541, 1995.
- [16] B. B. Biswal, M. Mennes, X.-N. Zuo, S. Gohel, C. Kelly, S. M. Smith, C. F. Beckmann, J. S. Adelstein, R. L. Buckner, S. Colcombe, A.-M. Dogonowski, M. Ernst, D. Fair, M. Hampson, M. J. Hoptman, J. S. Hyde, V. J. Kiviniemi, R. Kotter, S.-J. Li, C.-P. Lin, M. J. Lowe, C. Mackay, D. J. Madden, K. H. Madsen, D. S. Margulies, H. S. Mayberg, K. McMahon, C. S. Monk, S. H. Mostofsky, B. J. Nagel, J. J. Pekar, S. J. Peltier, S. E. Petersen, V. Riedl, S. A. R. B. Rombouts, B. Rypma, B. L. Schlaggar, S. Schmidt, R. D. Seidler, G. J. Siegle, C. Sorg, G.-J. Teng, J. Veijola, A. Villringer, M. Walter, L. Wang, X.-C. Weng, S. Whitfield-Gabrieli, P. Williamson, C. Windischberger, Y.-F. Zang, H.-Y. Zhang, F. X. Castellanos, and M. P. Milham. Toward discovery science of human brain function. *Proceedings of the National Academy of Sciences of the United States of America*, 107(10):4734–4739, 2010.
- [17] H. Blalock. *Social Statistics*. NY: McGraw-Hill, 1972.
- [18] M. Boly, C. Phillips, L. Tshibanda, A. Vanhaudenhuyse, M. Schabus, T. T. Dang-Vu, G. Moonen, R. Hustinx, P. Maquet, and S. Laureys. Intrinsic brain activity in altered states of consciousness: How conscious is the default mode of brain function? In *Annals of the New York Academy of Sciences*, pages 119–129. Blackwell Publishing, 2008.
- [19] M. Breakspear, J. R. Terry, and K. J. Friston. Modulation of excitatory synaptic coupling facilitates synchronization and complex dynamics in a biophysical model of neuronal dynamics. *Network-computation in neural systems*, 14(4):703–732, NOV 2003.
- [20] M. Brett, K. Christoff, R. Cusack, and J. Lancaster. Using the Talairach atlas with the MNI template. *NeuroImage*, 13(6):S85–S85, 2001.

- [21] J. Britz, D. van de Ville, and C. M. Michel. BOLD correlates of EEG topography reveal rapid resting-state network dynamics. *NeuroImage*, In Press, Uncorrected Proof, 2010.
- [22] M. J. Brookes, C. M. Stevenson, G. R. Barnes, A. Hillebrand, M. I. G. Simpson, S. T. Francis, and P. G. Morris. Beamformer reconstruction of correlated sources using a modified source model. *NeuroImage*, 34(4):1454–1465, 2007.
- [23] S. J. Broyd, C. Demanuele, S. Debener, S. K. Helps, C. J. James, and E. J. S. Sonuga-Barke. Default-mode brain dysfunction in mental disorders: A systematic review. *Neuroscience and Biobehavioral Reviews*, 33(3):279–296, 2009.
- [24] G. Buzsaki. *Rhythms of the Brain*. Oxford University Press, USA, 1st edition, August 2006.
- [25] G. Buzsaki, R. G. Bickford, D. M. Armstrong, G. Ponomareff, K. S. Chen, R. Ruiz, L. J. Thal, and F. H. Gage. Electric-activity in the neocortex of freely moving young and aged rats. *Neuroscience*, 26(3):735–744, 1988.
- [26] N. A. Carlson. *Foundations of Physiological Psychology*. Simon & Schuster, Needham Heights, Massachusetts, 1992.
- [27] C. Chang and G. H. Glover. Effects of model-based physiological noise correction on default mode network anti-correlations and correlations. *NeuroImage*, 47(4):1448–1459, 2009.
- [28] C. Chang and G. H. Glover. Time-frequency dynamics of resting-state brain connectivity measured with fMRI. *NeuroImage*, 50(1):81 – 98, 2010.
- [29] D. R. Chialvo. The brain near the edge. In Garrido, PL and Marro, J and Torres, JJ, editor, *Cooperative Behavior in Neural Systems*, volume 887 of *AIP Conference Proceedings*, pages 1–12. Amer Inst Physics, 2007. 9th Granada Seminar, Granada, SPAIN, SEP 11-15, 2006.
- [30] G. C. Chow. Tests of equality between sets of coefficients in 2 linear regressions. *Econometrica*, 28(3):591–605, 1960.
- [31] K. Christoff, A. M. Gordon, J. Smallwood, R. Smith, and J. W. Schooler. Experience sampling during fMRI reveals default network and executive system contributions to mind wandering. *Proceedings of the National Academy of Sciences of the United States of America*, 106(21):8719–8724, 2009.
- [32] P. S. Churchland and T. J. Sejnowski. Perspectives on cognitive neuroscience. *Science*, 242(4879):741–745, 1988.
- [33] D. M. Cole, S. M. Smith, and C. F. Beckmann. Advances and pitfalls in the analysis and interpretation of resting-state fMRI data. *Frontiers in Systems Neuroscience*, 4, 2010.
- [34] P. Comon. Independent component analysis, a new concept? *Signal Processing*, 36(3):287–314, 1994.

- [35] A. Compte, M. Sanchez-Vives, D. McCormick, and X. Wang. Cellular and network mechanisms of slow oscillatory activity (≈ 1 Hz) and wave propagations in a cortical network model. *Journal of Neurophysiology*, 89(5):2707–2725, 2003.
- [36] D. Cordes, V. M. Haughton, K. Arfanakis, G. J. Wendt, P. A. Turski, C. H. Moritz, M. A. Quigley, and M. E. Meyerand. Mapping functionally related regions of brain with functional connectivity MR imaging. *American Journal of Neuroradiology*, 21(9):1636–1644, 2000.
- [37] P. J. Davis. *Circulant matrices*. Wiley, New York, 1979.
- [38] M. de Luca, C. F. Beckmann, N. De Stefano, P. M. Matthews, and S. M. Smith. fMRI resting state networks define distinct modes of long-distance interactions in the human brain. *NeuroImage*, 29(4):1359–1367, 2006.
- [39] J. C. de Munck, S. I. Goncalves, R. Mammoliti, R. M. Heethaar, and F. H. L. da Silva. Interactions between different EEG frequency bands and their effect on alpha-fMRI correlations. *NeuroImage*, 47(1):69–76, 2009.
- [40] J. A. de Zwart, P. van Gelderen, J. M. Jansma, M. Fukunaga, M. Bianciardi, and J. H. Duyn. Hemodynamic nonlinearities affect BOLD fMRI response timing and amplitude. *NeuroImage*, 47(4):1649 – 1658, 2009.
- [41] R. H. R. Deckers, P. van Gelderen, M. Ries, O. Barret, J. H. Duyn, V. N. Ikonomidou, M. Fukunaga, G. H. Glover, and J. A. de Zwart. An adaptive filter for suppression of cardiac and respiratory noise in MRI time series data. *NeuroImage*, 33(4):1072–1081, 2006.
- [42] G. Deco, V. Jirsa, A. R. McIntosh, O. Sporns, and R. Koetter. Key role of coupling, delay, and noise in resting brain fluctuations. *Proceedings of the National Academy of Sciences of the United States of America*, 106(25):10302–10307, 2009.
- [43] A. Delorme and S. Makeig. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, 134(1):9–21, 2004.
- [44] G. Deshpande, S. LaConte, S. Peltier, and X. P. Hu. Connectivity analysis of human functional MRI data: From linear to nonlinear and static to dynamic. *Medical Imaging and Augmented Reality*, 4091:17–24, 2006.
- [45] R. A. Dineen, J. Vilisaar, J. Hlinka, C. M. Bradshaw, P. S. Morgan, C. S. Constantinescu, and D. P. Auer. Disconnection as a mechanism for cognitive dysfunction in multiple sclerosis. *Brain*, 132(1):239–249, 2009.
- [46] B. Ermentrout. *Simulating, Analyzing, and Animating Dynamical Systems: A Guide To Xppaut for Researchers and Students*. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2002.
- [47] G. B. Ermentrout. The behavior of rings of coupled oscillators. *Journal of Mathematical Biology*, 23(1):5574, 1985.

- [48] A. C. Evans, D. L. Collins, S. R. Mills, E. D. Brown, R. L. Kelly, and T. M. Peters. 3D statistical neuroanatomical models from 305 MRI volumes. In *Proc. IEEE-Nuclear Science Symposium and Medical Imaging Conference*, pages 1813–1817, 1993.
- [49] M. D. Fox, M. Corbetta, A. Z. Snyder, J. L. Vincent, and M. E. Raichle. Spontaneous neuronal activity distinguishes human dorsal and ventral attention systems. *Proceedings of the National Academy of Sciences of the United States of America*, 103(26):10046–10051, 2006.
- [50] M. D. Fox and M. E. Raichle. Spontaneous fluctuations in brain activity observed with functional magnetic resonance imaging. *Nature Reviews Neuroscience*, 8(9):700–711, 2007.
- [51] M. D. Fox, A. Z. Snyder, J. L. Vincent, M. Corbetta, D. C. Van Essen, and M. E. Raichle. The human brain is intrinsically organized into dynamic, anticorrelated functional networks. *Proceedings of the National Academy of Sciences of the United States of America*, 102(27):9673–9678, 2005.
- [52] M. D. Fox, D. Zhang, A. Z. Snyder, and M. E. Raichle. The Global Signal and Observed Anticorrelated Resting State Brain Networks. *Journal of Neurophysiology*, 101(6):3270–3283, 2009.
- [53] P. Fransson. Spontaneous low-frequency BOLD signal fluctuations: An fMRI investigation of the resting-state default mode of brain function hypothesis. *Human Brain Mapping*, 26(1):15–29, 2005.
- [54] P. Fransson. How default is the default mode of brain function? Further evidence from intrinsic BOLD signal fluctuations. *Neuropsychologia*, 44(14):2836–2845, 2006.
- [55] T. Freund, I. Katona, and D. Piomelli. Role of endogenous cannabinoids in synaptic signaling. *Physiological Reviews*, 83(3):1017–1066, 2003.
- [56] K. J. Friston. Functional and effective connectivity in neuroimaging: A synthesis. *Human Brain Mapping*, 2(1-2):56–78, 1994.
- [57] K. J. Friston. Causal modelling and brain connectivity in functional magnetic resonance imaging. *PLoS Biology*, 7(2):220–225, 2009.
- [58] K. J. Friston, J. Ashburner, S. J. Kiebel, T. E. Nichols, and W. D. Penny, editors. *Statistical Parametric Mapping: The Analysis of Functional Brain Images*. Academic Press, 2007.
- [59] K. J. Friston, C. D. Frith, P. F. Liddle, and R. S. J. Frackowiak. Functional Connectivity - The Principal-Component Analysis Of Large (PET) Data Sets. *Journal Of Cerebral Blood Flow And Metabolism*, 13(1):5–14, 1993.
- [60] K. J. Friston, L. Harrison, and W. Penny. Dynamic causal modelling. *NeuroImage*, 19(4):1273–1302, 2003.
- [61] K. J. Friston, S. Williams, R. Howard, R. S. J. Frackowiak, and R. Turner. Movement-related effects in fMRI time-series. *Magnetic Resonance in Medicine*, 35:346–355, 1996.

- [62] M. Fukunaga, S. G. Horovitz, P. van Gelderen, J. A. de Zwart, J. M. Jansma, V. N. Ikonomidou, R. X. Chu, R. H. R. Deckers, D. A. Leopold, and J. H. Duyn. Large-amplitude, spatially correlated fluctuations in BOLD fMRI signals during extended rest and early sleep stages. *Magnetic Resonance Imaging*, 24(8):979–992, 2006.
- [63] A. Ghosh, Y. Rho, A. R. McIntosh, R. Ktner, and V. K. Jirsa. Noise during rest enables the exploration of the brain’s dynamic repertoire. *PLoS Computational Biology*, 4(10):e1000196, 2008.
- [64] A. Gibson and H. Dehghani. Diffuse optical imaging. *Philosophical Transactions of the Royal Society A*, 367(1900):3055–3072, 2009.
- [65] S. J. Gilbert, I. Dumontheil, J. S. Simons, C. D. Frith, and P. W. Burgess. Comment on ”wandering minds: The default network and stimulus-independent thought”. *Science*, 317(5834):43b–, 2007.
- [66] F. Giove, T. Gili, V. Iacovella, E. Macaluso, and B. Maraviglia. Images-based suppression of unwanted global signals in resting-state functional connectivity studies. *Magnetic Resonance Imaging*, 27(8):1058–1064, 2009.
- [67] G. H. Glover, T. Q. Li, and D. Ress. Image-based method for retrospective correction of physiological motion effects in fMRI: RETROICOR. *Magnetic Resonance in Medicine*, 44(1):162–167, 2000.
- [68] R. Goldman, J. Stern, J. Engel, and M. Cohen. Simultaneous EEG and fMRI of the alpha rhythm. *Neuroreport*, 13(18):2487–2492, 2002.
- [69] D. Goodman and R. Brette. Brian: a simulator for spiking neural networks in Python. *Frontiers in Neuroinformatics*, 2, 2008.
- [70] C. Granger. Investigating causal relations by econometric model and cross-spectral methods. *Econometrica*, 37(3):414–&, 1969.
- [71] M. D. Greicius, B. H. Flores, V. Menon, G. H. Glover, H. B. Solvason, H. Kenna, A. L. Reiss, and A. F. Schatzberg. Resting-state functional connectivity in major depression: Abnormally increased contributions from subgenual cingulate cortex and thalamus. *Biological Psychiatry*, 62(5):429–437, 2007.
- [72] M. D. Greicius, V. Kiviniemi, O. Tervonen, V. Vainionpaa, S. Alahuhta, A. L. Reiss, and V. Menon. Persistent default-mode network connectivity during light sedation. *Human Brain Mapping*, 29(7):839–847, 2008.
- [73] M. D. Greicius, G. Srivastava, A. L. Reiss, and V. Menon. Default-mode network activity distinguishes Alzheimer’s disease from healthy aging: Evidence from functional MRI. *Proceedings of the National Academy of Sciences of the United States of America*, 101(13):4637–4642, 2004.
- [74] J. Gross, J. Kujala, M. Hamalainen, L. Timmermann, A. Schnitzler, and R. Salmelin. Dynamic imaging of coherent sources: Studying neural interactions in the human brain. *Proceedings of the National Academy of Sciences of the United States of America*, 98(2):694–699, 2001.

- [75] J. Guckenheimer and R. A. Oliva. Chaos in the Hodgkin–Huxley Model. *SIAM Journal on Applied Dynamical Systems*, 1:105–114, 2002.
- [76] M. E. Haacke, R. W. Brown, M. R. Thompson, R. Venkatesan, M. E. Haacke, R. W. Brown, M. R. Thompson, and R. Venkatesan. *Magnetic Resonance Imaging: Physical Principles and Sequence Design*. Wiley-Liss, 1st edition, June 1999.
- [77] S. Han, C. Kurrer, and Y. Kuramoto. Dephasing and bursting in coupled neural oscillators. *Physical Review Letters*, 75(17):3190–3193, 1995.
- [78] J. Hlinka, C. Alexakis, A. Diukova, P. F. Liddle, and D. P. Auer. Slow EEG pattern predicts reduced intrinsic functional connectivity in the default mode network: An inter-subject analysis. *NeuroImage*, 53:239–246, 2010.
- [79] J. Hlinka, C. Alexakis, J. G. Hardman, Q. Siddiqui, and D. P. Auer. Is sedation-induced BOLD fMRI low-frequency fluctuation increase mediated by increased motion? *Magnetic Resonance Materials in Physics, Biology and Medicine*, Epub ahead of print, 2010.
- [80] J. Hlinka and S. Coombes. Depolarization Induced Suppression of Excitation and the Emergence of Ultraslow Rhythms in Neural Networks. *Physical Review Letters*, 104(6), 2010.
- [81] J. Hlinka, M. Palus, M. Vejmelka, D. Mantini, and M. Corbetta. Functional connectivity in resting-state fMRI: Is linear correlation sufficient? *NeuroImage*, 2010, In Press.
- [82] C. J. Honey, R. Kotter, M. Breakspear, and O. Sporns. Network structure of cerebral cortex shapes functional connectivity on multiple time scales. *Proceedings of the National Academy of Sciences of the United States of America*, 104(24):10240–10245, 2007.
- [83] C. J. Honey and O. Sporns. Dynamical consequences of lesions in cortical networks. *Human Brain Mapping*, 29(7):802–809, 2008.
- [84] C. J. Honey, O. Sporns, L. Cammoun, X. Gigandet, J. P. Thiran, R. Meuli, and P. Hagmann. Predicting human resting-state functional connectivity from structural connectivity. *Proceedings of the National Academy of Sciences of the United States of America*, 106(6):2035–2040, 2009.
- [85] F. C. Hoppensteadt and E. M. Izhikevich. *Weakly connected neural networks*. Springer-Verlag New York, Inc., Secaucus, NJ, USA, 1997.
- [86] S. G. Horovitz, A. R. Braun, W. S. Carr, D. Picchioni, T. J. Balkin, M. Fukunaga, and J. H. Duyn. Decoupling of the brain’s default mode network during deep sleep. *Proceedings of the National Academy of Sciences of the United States of America*, 106(27):11376–11381, 2009.
- [87] S. G. Horovitz, M. Fukunaga, J. A. de Zwart, P. van Gelderen, S. C. Fulton, T. J. Balkin, and J. H. Duyn. Low frequency BOLD fluctuations during resting wakefulness and light sleep: A simultaneous EEG-fMRI study. *Human Brain Mapping*, 29(6):671–682, 2008.

- [88] A. G. Hudetz, J. J. Smith, J. G. Lee, Z. J. Bosnjak, and J. P. Kampine. Modification of Cerebral Laser-Doppler Flow Oscillations by Halothane, Pco(2), and Nitric-Oxide Synthase Blockade. *American Journal of Physiology-Heart and Circulatory Physiology*, 38(1):H114–H120, 1995.
- [89] W. G. Hundley, G. J. Renaldo, J. E. Levasseur, and H. A. Kontos. Vasomotion in cerebral microcirculation of awake rabbits. *American Journal of Physiology*, 254(1):H67–H71, 1988.
- [90] K. Jann, T. Dierks, C. Boesch, M. Kottlow, W. Strik, and T. Koenig. BOLD correlates of EEG alpha phase-locking and the fMRI default mode network. *NeuroImage*, 45(3):903–916, 2009.
- [91] M. Jenkinson. Measuring Transformation Error by RMS Deviation. Technical report, Oxford, 1999.
- [92] V. Jirsa and A. McIntosh, editors. *Handbook of Brain Connectivity*. Springer, 2007.
- [93] M. Kaiser and C. C. Hilgetag. Optimal hierarchical modular topologies for producing limited sustained activation of neural networks. *Frontiers in Neuroinformatics*, 2010.
- [94] E. R. Kandel, J. H. Schwartz, and T. M. Jessell. *Principles of Neural Science*. McGraw-Hill Medical, 4th edition, July 2000.
- [95] R. Kingsley. *Concise Text of Neuroscience*. Lippincott Williams & Wilkins, Philadelphia, 1996.
- [96] V. Kiviniemi. Endogenous brain fluctuations and diagnostic imaging. *Human Brain Mapping*, 29(7):810–817, 2008.
- [97] V. Kiviniemi, J. Kantola, J. Jauhiainen, A. Hyvarinen, and O. Tervonen. Independent component analysis of nondeterministic fMRI signal sources. *NeuroImage*, 19(2, Part 1):253–260, 2003.
- [98] V. Kiviniemi, T. Starck, J. Remes, X. Long, J. Nikkinen, M. Haapea, J. Veijola, I. Moilanen, M. Isohanni, Y.-F. Zang, and O. Tervonen. Functional Segmentation of the Brain Cortex Using High Model Order Group PICA. *Human Brain Mapping*, 30(12):3865–3886, 2009.
- [99] V. J. Kiviniemi, H. Haanpaa, J. H. Kantola, J. Jauhiainen, W. Vainionpaa, S. Alahuhta, and O. Tervonen. Midazolam sedation increases fluctuation and synchrony of the resting brain BOLD signal. *Magnetic Resonance Imaging*, 23(4):531–537, 2005.
- [100] G. G. Knyazev, A. N. Savostyanov, and E. A. Levin. Alpha oscillations as a correlate of trait anxiety. *International Journal of Psychophysiology*, 53(2):147–160, 2004.
- [101] C. Koch and I. Segev, editors. *Methods in Neuronal Modeling: From Ions to Networks*. MIT Press, Cambridge, MA, USA, 1998.

- [102] R. Kotter. Online retrieval, processing, and visualization of primate connectivity data from the CoCoMac database. *Neuroinformatics*, 2(2):127–144, 2004.
- [103] Y. T. Kwak. Quantitative EEG findings in different stages of Alzheimer’s disease. *Journal of Clinical Neurophysiology*, 23(5):456–461, 2006.
- [104] P. J. Lahaye, J. B. Poline, G. Flandin, S. Dodel, and L. Garnero. Functional connectivity: studying nonlinear, delayed interactions between BOLD signals. *NeuroImage*, 20(2):962–974, 2003.
- [105] E. Langford, N. Schwartzman, and M. Owens. Is the property of being positively correlated transitive? *American Statistician*, 55(4):322–325, 2001.
- [106] L. J. Larson-Prior, J. M. Zempel, T. S. Nolan, F. W. Prior, A. Z. Snyder, and M. E. Raichle. Cortical network functional connectivity in the descent to sleep. *Proceedings of the National Academy of Sciences of the United States of America*, 106(11):4489–4494, 2009.
- [107] H. Laufs, J. L. Holt, R. Elfont, M. Krams, J. S. Paul, K. Krakow, and A. Kleinschmidt. Where the BOLD signal goes when alpha EEG leaves. *NeuroImage*, 31(4):1408–1418, 2006.
- [108] H. Laufs, A. Kleinschmidt, A. Beyerle, E. Eger, A. Salek-Haddadi, C. Preibisch, and K. Krakow. EEG-correlated fMRI of human alpha activity. *NeuroImage*, 19(4):1463–1476, 2003.
- [109] H. Laufs, K. Krakow, P. Sterzer, E. Eger, A. Beyerle, A. Salek-Haddadi, and A. Kleinschmidt. Electroencephalographic signatures of attentional and cognitive default modes in spontaneous brain activity fluctuations at rest. *Proceedings of the National Academy of Sciences of the United States of America*, 100(19):11053–11058, 2003.
- [110] D. A. Leopold, Y. Murayama, and N. K. Logothetis. Very slow activity fluctuations in monkey visual cortex: Implications for functional brain imaging. *Cerebral Cortex*, 13(4):422–433, 2003.
- [111] K. Linkenkaer-Hansen, V. Nikouline, J. Palva, and R. Ilmoniemi. Long-range temporal correlations and scaling behavior in human brain oscillations. *Journal of Neuroscience*, 21(4):1370–1377, 2001.
- [112] X. Liu, X. Zhu, Y. Zhang, and W. Chen. Insights into the origin of spontaneous coherent BOLD fluctuations in a resting rat brain under varied isoflurane anesthesia depth. In *Proceedings 17th Scientific Meeting, International Society for Magnetic Resonance in Medicine*, page 123, 2009.
- [113] N. K. Logothetis and B. A. Wandell. Interpreting the BOLD signal. *Annual Review of Physiology*, 66:735–769, 2004.
- [114] L. S. Ma, B. Q. Wang, X. Y. Chen, and J. H. Xiong. Detecting functional connectivity in the resting brain: a comparison between ICA and CCA. *Magnetic Resonance Imaging*, 25(1):47–56, 2007.

- [115] D. P. MacKinnon, A. J. Fairchild, and M. S. Fritz. Mediation analysis. *Annual Review of Psychology*, 58:593–614, 2007.
- [116] D. P. MacKinnon, C. M. Lockwood, J. M. Hoffman, S. G. West, and V. Sheets. A comparison of methods to test mediation and other intervening variable effects. *Psychological Methods*, 7(1):83–104, 2002.
- [117] P. Mansfield. Multi-planar image-formation using NMR spin echoes. *Journal of Physics C*, 10(3):L55–L58, 1977.
- [118] D. Mantini, M. G. Perrucci, C. Del Gratta, G. L. Romani, and M. Corbetta. Electrophysiological signatures of resting state networks in the human brain. *Proceedings of the National Academy of Sciences of the United States of America*, 104(32):13170–13175, 2007.
- [119] R. Martuzzi, R. Ramani, M. Qiu, N. Rajeevan, and R. T. Constable. Functional connectivity and alterations in baseline brain state in humans. *NeuroImage*, 49(1):823 – 834, 2010.
- [120] M. F. Mason, M. I. Norton, J. D. Van Horn, D. M. Wegner, S. T. Grafton, and C. N. Macrae. Wandering minds: The default network and stimulus-independent thought. *Science*, 315(5810):393–395, 2007.
- [121] V. Maxim, L. Sendur, J. Fadili, J. Suckling, R. Gould, R. Howard, and E. Bullmore. Fractional Gaussian noise, functional MRI and Alzheimer’s disease. *NeuroImage*, 25(1):141–158, 2005.
- [122] B. Mazoyer, L. Zago, E. Mellet, S. Bricogne, O. Etard, O. Houde, F. Crivello, M. Joliot, L. Petit, and N. Tzourio-Mazoyer. Cortical networks for working memory and executive functions sustain the conscious resting state in man. *Brain Research Bulletin*, 54(3):287–298, 2001.
- [123] M. McAvoy, L. Larson-Prior, T. S. Nolan, S. N. Vaishnavi, M. E. Raichle, and G. d’Avossa. Resting states affect spontaneous BOLD oscillations in sensory and paralimbic cortex. *Journal of Neurophysiology*, 100(2):922–931, 2008.
- [124] W. S. McCulloch and W. H. Pitts. A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics*, 5:115–133, 1943.
- [125] M. McKeown, S. Makeig, G. Brown, T. Jung, S. Kindermann, A. Bell, and T. Sejnowski. Analysis of fMRI data by blind separation into independent spatial components. *Human Brain Mapping*, 6(3):160–188, 1998.
- [126] T. Meindl, S. Teipel, R. Elmouden, S. Mueller, W. Koch, O. Dietrich, U. Coates, M. Reiser, and C. Glaser. Test-Retest Reproducibility of the Default-Mode Network in Healthy Individuals. *Human Brain Mapping*, 31(2):237–246, 2010.
- [127] K. J. Miller, K. E. Weaver, and J. G. Ojemann. Direct electrophysiological measurement of human default network areas. *Proceedings of the National Academy of Sciences of the United States of America*, 106(29):12174–12177, 2009.

- [128] M. Moosmann, P. Ritter, I. Krastel, A. Brink, S. Thees, F. Blankenburg, B. Taskin, H. Obrig, and A. Villringer. Correlates of alpha rhythm in functional magnetic resonance imaging and near infrared spectroscopy. *NeuroImage*, 20(1):145–158, 2003.
- [129] A. M. Morcom and P. C. Fletcher. Does the brain have a baseline? Why we should be resisting a rest. *NeuroImage*, 37(4):1073–1082, 2007.
- [130] C. Morris and H. Lecar. Voltage oscillations in the barnacle giant muscle-fiber. *Biophysical Journal*, 35(1):193–213, 1981.
- [131] J. Müller-Gerking, J. Martinerie, S. Neuenschwander, L. Pezard, B. Renault, and F. Varela. Detecting non-linearities in neuro-electrical signals: A study of synchronous local field potentials. *Physica D: Nonlinear Phenomena*, 94(1-2):65–91, 1996.
- [132] K. Murphy, R. M. Birn, D. A. Handwerker, T. B. Jones, and P. A. Bandettini. The impact of global signal regression on resting state correlations: Are anti-correlated networks introduced? *NeuroImage*, 44(3):893–905, 2009.
- [133] F. Musso, J. Brinkmeyer, A. Mobascher, T. Warbrick, and G. Winterer. Spontaneous brain activity and EEG microstates. A novel EEG/fMRI analysis approach to explore resting-state networks. *NeuroImage*, In Press, Corrected Proof:–, 2010.
- [134] R. Nagarajan. Surrogate testing of linear feedback processes with non-gaussian innovations. *Physica A*, 366(1):530–538, 2006.
- [135] Y. Nir, R. Mukamel, I. Dinstein, E. Privman, M. Harel, L. Fisch, H. Gelbard-Sagiv, S. Kipervasser, F. Andelman, M. Y. Neufeld, U. Kramer, A. Arieli, I. Fried, and R. Malach. Interhemispheric correlations of slow spontaneous neuronal fluctuations revealed in human sensory cortex. *Nature Neuroscience*, 11(9):1100–1108, 2008.
- [136] S. Ogawa, T. M. Lee, A. R. Kay, and D. W. Tank. Brain magnetic-resonance-imaging with contrast dependent on blood oxygenation. *Proceedings of the National Academy of Sciences of the United States of America*, 87(24):9868–9872, 1990.
- [137] T. Ohno-Shosaku, T. Maejima, and M. Kano. Endogenous cannabinoids mediate retrograde signals from depolarized postsynaptic neurons to presynaptic terminals. *Neuron*, 29(3):729–738, 2001.
- [138] M. Palus. Detecting phase synchronization in noisy systems. *Physics Letters A*, 235(4):341–351, 1997.
- [139] M. Palus. From nonlinearity to causality: statistical testing and inference of physical mechanisms underlying complex dynamics. *Contemporary Physics*, 48(6):307–348, 2007.
- [140] M. Palus, V. Albrecht, and I. Dvorak. Information theoretic test for non-linearity in time series. *Physics Letters A*, 175(3-4):203–209, 1993.

- [141] M. Palus and M. Vejmelka. Directionality of coupling from bivariate time series: How to avoid false causalities and missed connections. *Physical Review E (Statistical, Nonlinear, and Soft Matter Physics)*, 75(5, Part 2), 2007.
- [142] S. J. Peltier, C. Kerssens, S. B. Hamann, P. S. Sebel, M. Byas-Smith, and X. P. Hu. Functional connectivity changes with concentration of sevoflurane anesthesia. *Neuroreport*, 16(3):285–288, 2005.
- [143] K. J. Preacher and A. F. Hayes. Spss and sas procedures for estimating indirect effects in simple mediation models. *Behavior Research Methods Instruments & Computers*, 36(4):717–731, 2004.
- [144] K. J. Preacher and A. F. Hayes. Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior Research Methods*, 40(3):879–891, 2008.
- [145] D. Prichard and J. Theiler. Generating surrogate data for time series with several simultaneously measured variables. *Physical Review Letters*, 73(7):951, 1994.
- [146] M. E. Raichle. A Paradigm Shift in Functional Brain Imaging. *Journal of Neuroscience*, 29(41):12729–12734, 2009.
- [147] M. E. Raichle and D. A. Gusnard. Intrinsic brain activity sets the stage for expression of motivated behavior. *Journal of Comparative Neurology*, 493(1):167–176, 2005. Neuroscience Conference on Anatomy of the Soul, Ameland, Netherlands, May 19-24, 2005.
- [148] M. E. Raichle, A. M. MacLeod, A. Z. Snyder, W. J. Powers, D. A. Gusnard, and G. L. Shulman. A default mode of brain function. *Proceedings of the National Academy of Sciences of the United States of America*, 98(2):676–682, 2001.
- [149] M. E. Raichle and A. Z. Snyder. A default mode of brain function: A brief history of an evolving idea. *NeuroImage*, 37(4):1083–1090, 2007.
- [150] M. A. E. Ramsay, T. M. Savege, B. R. J. Simpson, and R. Goodwin. Controlled sedation with alphaxalone-alphadolone. *British Medical Journal*, 2(5920):656–659, 1974.
- [151] M. Razavi, B. Eaton, S. Paradiso, M. Mina, A. G. Hudetz, and L. Bolinger. Source of low-frequency fluctuations in functional mri signal. *Journal of Magnetic Resonance Imaging*, 27(4):891–897, 2008.
- [152] D. Robbe, S. M. Montgomery, A. Thome, P. E. Rueda-Orozco, B. L. McNaughton, and G. Buzsaki. Cannabinoids reveal importance of spike timing coordination in hippocampal function. *Nature Neuroscience*, 9(12):1526–1533, 2006.
- [153] A. Roebroeck, E. Formisano, and R. Goebel. The identification of interacting networks in the brain using fMRI: Model selection, causality and deconvolution. *NeuroImage*, In Press, Corrected Proof:–, 2009.

- [154] S. Rombouts, R. Keunen, and C. STAM. Investigation of nonlinear structure in multichannel EEG. *Physics Letters A*, 202(5-6):352–358, 1995.
- [155] P. Sämann, R. Wehrle, V. Spoormaker, D. Höhn, H. Peters, F. Holsboer, and M. Czisch. Development of the brain default mode network from wakefulness into slow wave sleep. In *Proceedings 17th Scientific Meeting, International Society for Magnetic Resonance in Medicine*, page 126, 2009.
- [156] R. Scheeringa, M. C. M. Bastiaansen, K. M. Petersson, R. Oostenveld, D. G. Norris, and P. Hagoort. Frontal theta EEG activity correlates negatively with the default mode network in resting state. *International Journal of Psychophysiology*, 67(3):242–251, 2008.
- [157] T. Schreiber. Measuring information transfer. *Physical Review Letters*, 85(2):461–464, 2000.
- [158] D. Schwender, M. Daunerer, S. Klasing, U. Finsterer, and K. Peter. Power spectral analysis of the electroencephalogram during increasing end-expiratory concentrations of isoflurane, desflurane and sevoflurane. *Anaesthesia*, 53(4):335–342, 1998.
- [159] W. W. Seeley, V. Menon, A. F. Schatzberg, J. Keller, G. H. Glover, H. Kenna, A. L. Reiss, and M. D. Greicius. Dissociable intrinsic connectivity networks for salience processing and executive control. *Journal of Neuroscience*, 27(9):2349–2356, 2007.
- [160] Z. Shehzad, A. M. C. Kelly, P. T. Reiss, D. G. Gee, K. Gotimer, L. Q. Uddin, S. H. Lee, D. S. Margulies, A. K. Roy, B. B. Biswal, E. Petkova, F. X. Castellanos, and M. P. Milham. The Resting Brain: Unconstrained yet Reliable. *Cerebral Cortex*, 19(10):2209–2229, 2009.
- [161] G. L. Shulman, J. A. Fiez, M. Corbetta, R. L. Buckner, F. M. Miezin, M. E. Raichle, and S. E. Petersen. Common blood flow changes across visual tasks. II: Decreases in cerebral cortex. *Journal of Cognitive Neuroscience*, 9(5):648–663, 1997.
- [162] J. W. Sleigh and J. Donovan. Comparison of bispectral index, 95approximate entropy of the EEG, with changes in heart rate variability during induction of general anaesthesia. *British Journal of Anaesthesia*, 82(5):666–671, 1999.
- [163] D. Smirnov and R. Andrzejak. Detection of weak directional coupling: Phase-dynamics approach versus state-space approach. *Physical Review E (Statistical, Nonlinear, and Soft Matter Physics)*, 71(3, Part 2), 2005.
- [164] A. M. Smith, B. K. Lewis, U. E. Ruttimann, F. Q. Ye, T. M. Sinnwell, Y. H. Yang, J. H. Duyn, and J. A. Frank. Investigation of low frequency drift in fMRI signal. *NeuroImage*, 9(5):526–533, 1999.
- [165] S. M. Smith, P. T. Fox, K. L. Miller, D. C. Glahn, P. M. Fox, C. E. Mackay, N. Filippini, K. E. Watkins, R. Toro, A. R. Laird, and C. F. Beckmann. Correspondence of the brain’s functional architecture during activation and rest. *Proceedings of the National Academy of Sciences of the United States of America*, 106(31):13040–13045, 2009.

- [166] S. M. Smith, M. Jenkinson, M. W. Woolrich, C. F. Beckmann, T. E. J. Behrens, H. Johansen-Berg, P. R. Bannister, M. De Luca, I. Drobnjak, D. E. Flitney, R. K. Niazy, J. Saunders, J. Vickers, Y. Y. Zhang, N. De Stefano, J. M. Brady, and P. M. Matthews. Advances in functional and structural MR image analysis and implementation as FSL. *NeuroImage*, 23:S208–S219, 2004.
- [167] G. Srivastava, S. Crottaz-Herbette, K. M. Lau, G. H. Glover, and V. Menon. ICA-based procedures for removing ballistocardiogram artifacts from EEG data acquired in the MRI scanner. *NeuroImage*, 24(1):50–60, 2005.
- [168] C. Stam. Nonlinear dynamical analysis of EEG and MEG: Review of an emerging field. *Clinical Neurophysiology*, 116(10):2266–2301, 2005.
- [169] I. StatSoft. Electronic statistics textbook, 2006.
- [170] M. Steriade. Grouping of brain rhythms in corticothalamic systems. *Neuroscience*, 137(4):1087–1106, 2006.
- [171] A. Straiker and K. Mackie. Depolarization-induced suppression of excitation in murine autaptic hippocampal neurones. *Journal of Physiology*, 569(2):501–517, 2005.
- [172] J. R. Terry and M. Breakspear. An improved algorithm for the detection of dynamical interdependence in bivariate time-series. *Biological Cybernetics*, 88(2):129–136, 2003.
- [173] B. Thirion, S. Dodel, and J. B. Poline. Detection of signal synchronizations in resting-state fMRI datasets. *NeuroImage*, 29(1):321–327, 2006.
- [174] M. van den Heuvel, R. Mandl, and H. Hulshoff Pol. Normalized cut group clustering of resting-state fmri data. *PLoS ONE*, 3(4):e2001, 2008.
- [175] K. R. A. van Dijk, T. Hedden, A. Venkataraman, K. C. Evans, S. W. Lazar, and R. L. Buckner. Intrinsic functional connectivity as a tool for human connectomics: Theory, properties, and optimization. *Journal of Neurophysiology*, 103(1):297–321, 2010.
- [176] J. L. Vincent, G. H. Patel, M. D. Fox, A. Z. Snyder, J. T. Baker, D. C. Van Essen, J. M. Zempel, L. H. Snyder, M. Corbetta, and M. E. Raichle. Intrinsic functional architecture in the anaesthetized monkey brain. *Nature*, 447(7140):83–U4, 2007.
- [177] H. von Helmholtz. Ueber einige Gesetze der Vertheilung elektrischer Ströme in körperlichen Leitern, mit Anwendung auf die thierisch-electrischen Versuche. *Annalen Der Physik Und Chemie*, 89:354–357, 1853.
- [178] H. R. Wilson and J. D. Cowan. Excitatory and inhibitory interactions in localized populations of model neurons. *Biophysical Journal*, 12:1–24, 1972.
- [179] R. Wilson and R. Nicoll. Neuroscience - Endocannabinoid signaling in the brain. *Science*, 296(5568):678–682, 2002.

- [180] X. P. Xie, Z. T. Cao, and X. C. Weng. Spatiotemporal nonlinearity in resting-state fMRI of the human brain. *NeuroImage*, 40(4):1672–1685, 2008.
- [181] L. Zemanova, C. Zhou, and J. Kurths. Structural and functional clusters of complex brain networks. *Physica D: Nonlinear Phenomena*, 224(1-2):202–212, 2006.
- [182] Q.-H. Zou, C.-Z. Zhu, Y. Yang, X.-N. Zuo, X.-Y. Long, Q.-J. Cao, Y.-F. Wang, and Y.-F. Zang. An improved approach to detection of amplitude of low-frequency fluctuation (ALFF) for resting-state fMRI: Fractional ALFF. *Journal of Neuroscience Methods*, 172(1):137–141, 2008.
- [183] X.-N. Zuo, C. Kelly, J. S. Adelstein, D. F. Klein, F. X. Castellanos, and M. P. Milham. Reliable intrinsic connectivity networks: Test-retest evaluation using ICA and dual regression approach. *NeuroImage*, 49(3):2163–2177, 2010.