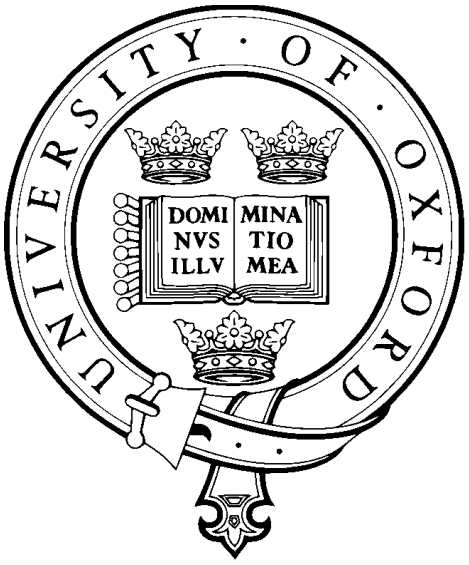
Investigating the role of APOE-ϵ4, a risk gene for Alzheimer's disease, on functional brain networks using magnetoencephalography



Henry Luckhoo

Trinity College

Centre For Doctoral Training In Healthcare Innovation
Oxford Centre for Human Brain Activity

D.Phil Thesis
Department of Engineering Science

Supervised by Dr Mark Woolrich

Submitted: September 25, 2013.

This thesis is submitted to the Department of Engineering Science, University of Oxford, in fulfilment of the requirements for the degree of Doctor of Philosophy

Investigating the role of APOE-ϵ4, a risk gene for Alzheimer's disease, on functional brain networks using magnetoencephalography

Henry Luckhoo

Trinity College

Abstract

Alzheimer's disease (AD) is developing into the single greatest healthcare challenge in the coming decades. The development of early and effective treatments that can prevent the pathological damage responsible for AD-related dementia is of utmost priority for healthcare authorities. The role of the APOE-ϵ4 genotype, which has been shown to increase an individual's risk of developing AD, is of central interest to this goal. Understanding the mechanism by which possession of this gene modulates brain function, leading to a predisposition towards AD is an active area of research. Functional connectivity (FC) is an excellent candidate for linking APOE-related differences in brain function to sites of AD pathology. Magnetoencephalography (MEG) is a neuroimaging tool that can provide a unique insight into the electrophysiology underpinning resting-state networks (RSNs) - whose dysfunction is postulated to lead to a predisposition to AD.

This thesis presents a range of methods for measuring functional connectivity in MEG data. We first develop a set of novel adaptations for preprocessing MEG data and performing source reconstruction using a beamformer (chapter 3). We then develop a range of analyses for measuring FC through correlations in the slow envelope oscillations of band-limited source-space MEG data (chapter 4). We investigate the optimum time scales for detecting FC. We then develop methods for extracting single networks (using seed-based correlation) and multiple networks (using ICA). We proceed to develop a group-statistical framework for detecting spatial differences in RSNs and present a preliminary finding for APOE-genotype-dependent differences in RSNs (chapter 5). We also develop a statistical framework for quantifying task-locked temporal differences in functional networks during task-positive experiments (chapter 6). Finally, we demonstrate a data-driven parcellation and network analysis pipeline that includes a novel correction for signal leakage between parcels. We use this framework to show evidence of stationary cross-frequency FC (chapter 7).

This thesis is submitted for the degree of Doctor of Philosophy
at the University of Oxford.
Michaelmas Term, 2013

Extended Abstract

Chapter 1: Connecting Alzheimer's disease, APOE-ϵ4 and resting-state networks: the role that MEG can play

We present the case for prioritising research into Alzheimer's disease (AD), highlighting that within the next few decades, AD threatens to overwhelm healthcare systems globally. We introduce the gene, APOE-ϵ4, which is a proven risk factor for AD. We compare the neuroimaging modalities capable of investigating the effect that APOE-ϵ4 has on brain function. Functional connectivity (FC) is introduced as a particularly promising research avenue for linking APOE-ϵ4, abnormal brain function in the resting-state networks of healthy individuals, and AD-related pathological damage in AD patients. Finally, we outline the novel insights that magnetoencephalography (MEG) can provide to this research area.

Chapter 2: An introduction to MEG

Here, we expand on what the signals MEG systems measure are, how MEG does this, and how we can analyse the data. We start by considering the neuronal mechanisms by which a measurable MEG signal is generated. We then provide a brief introduction into the instrumentation needed to measure these signals. We present the main approaches to analysing MEG, including the MEG *forward* and *inverse problems*. We compare the most popular MEG forward models currently available and then proceed to introduce methods for performing MEG source reconstruction. Here, we focus on the source reconstruction technique of beamforming as all of our functional connectivity analyses utilise beamformer-reconstructed MEG data. Finally, we contrast two major types of MEG system: the *Elekta Neuromag* and the *CTF 275-channel system*. Data from both of these systems are used throughout this thesis.

Chapter 3: Preprocessing for MEG analyses

In this chapter, we outline the MEG preprocessing framework that enables us to perform the functional connectivity analyses used throughout this thesis. We adopt a MEG preprocessing pipeline with three major novel contributions. The first is a procedure for applying *MaxFilter*^{TM} to *Elekta Neuromag* data that is robust to corruption from specific channel artefacts. The second is an implementation of independent component

analysis (ICA) for de-noising multi-sensor-type MEG data in sensor-space. The third is a set of modifications to the LCMV beamformer that enables source reconstruction of multi-sensor-type **and** *MaxFiltered* MEG data.

Chapter 4: Imaging functionally-connected networks in MEG

We present three sets of functional connectivity analyses for resting-state MEG. FC is defined as correlations in the slow *co-variation* oscillations of band-limited *carrier* time series of source-space neural activity. We initially investigate the optimum *carrier* and *co-variation* frequency window for measuring FC in the presence of signal leakage. We proceed to develop a seed-based correlation analysis for imaging single RSNs. We contrast several approaches for accounting for spurious FC introduced by signal leakage. We demonstrate our seed-based analysis by imaging the sensorimotor network. Finally, we present a framework for using ICA to decompose group-concatenated MEG envelope time courses into multiple RSNs. We demonstrate that this analysis works on MEG data acquired using a *CTF* system and using an *Elekta Neuromag* system (which represents the first successful extraction of RSNs from *Elekta* data).

Chapter 5: Investigating group differences in resting-state networks with MEG

In this chapter, we show, both theoretically and empirically, that beamformer *weights normalisation* is a major confound for multi-session statistics in MEG and must be accounted for accordingly. We go on to present a novel framework, based on dual regression, for evaluating differences in functional networks across different sub-groups. We demonstrate that MEG-adapted dual regression (MADR) can utilise **any** group average spatial basis set and determine sub-group differences relative to that average. We use MADR to show that carriers of the APOE-ϵ4 gene have significantly lower variance of oscillatory power associated with left and right frontoparietal networks, when compared with normal APOE-ϵ3 homozygotes.

Chapter 6: Analysis of functional networks in task-positive paradigms

In this chapter, we demonstrate that our functional connectivity analysis, based on performing temporal ICA on the slow-moving envelopes of band-limited neural activity, can be used to extract functionally-connected networks in task-positive MEG data. We present a *post hoc* analysis of the independent components using a mixed-effects General Linear Model (GLM). We show that our combined ICA/GLM analysis can provide advantages in task-positive MEG analysis which are unavailable to either technique when used in isolation. We apply our analysis to a *2-back* working memory experiment and show that it matches previous features associated with working memory function, outperforms a standard, voxel-wise GLM analysis and provides novel insights into working memory function in MEG.

Chapter 7: Parcellation and network analyses for MEG

In this chapter, we present a framework for estimating a data-driven parcellation for MEG using spatial ICA, and for performing a subsequent partial correlation network analysis to characterise the network structure between parcels. In particular, we focus on developing a novel method for removing the signal leakage contribution between a set of parcels that is compatible with multi-parcel network modelling methods such as partial correlation. We validate our analysis by estimating the standard and partial cross-frequency correlation matrices for the slow envelope oscillations of the band-limited neural activity of each parcel. We present preliminary evidence for the existence of stationary direct FC between the slow envelope oscillations of different frequency bands, both within and between parcels.

Chapter 8: Conclusions and future work

We conclude this thesis by considering the future directions in which the research presented here can be taken. We identify methods that can capture the dynamic nature of functional connectivity and methods that can combine information from multiple neuroimaging modalities as the most important areas of future research. We end by summarising this thesis as a set of methods that have been essential in linking the well-established field of fMRI functional connectivity to the much newer field of MEG functional connectivity.

Acknowledgements

I would like to thank the following institutions from the University of Oxford:

The Centre for Doctoral Training in Healthcare Innovation for funding this D.Phil under the RCUK Digital Economy Programme.

The Department of Engineering Science, the Department of Psychiatry and Trinity College for providing administrative and financial support.

I would also like to thank the Organisation for Human Brain Mapping, the Guarantors of Brain and the Wellcome Trust for providing funding for conference travel.

I am also greatly indebted to the following people without whose help this thesis would not be possible:

My supervisor, Dr Mark Woolrich for his outstanding supervision, guidance and mentorship over the last three years.

Prof. Matthew Brookes, who has been a fantastic collaborator and has been generous with his time, wisdom and resources.

Verena Heise and Dr Claire Mackay, who provided the *APOE* data set that forms one of the foundations of this work and without whom the core contributions of this D.Phil would not have been possible.

Dr Joanne Hale and Sofia Palazzo Corner for collecting the *CTF* data sets that form the other foundation of this thesis, using the resources at the Sir Peter Mansfield Magnetic Resonance Centre, University of Nottingham.

The analysis group of the Oxford Centre for Human Brain Activity (OHBA), including Adam Baker, Tyler Ferro, Sam Harrison, Giles Colclough, and Dr Hamid Mohseni.

The members of OHBA, including Prof. Kia Nobre, Dr Sven Braeutigam, Dr Mark Stokes, Dr Louisa Stacey, and George Wallis.

The analysis group at the Centre for Function Magnetic Resonance Imaging of the Brain (FMRIB), in particular Prof. Steve Smith and Dr Eugene Duff.

Jennifer Juillard-Maniece for her essential support throughout my D.Phil and her unrivalled German language skills.

My family, in particular my parents Keith and Jane Luckhoo, for everything they have given me.

Contents

List of Figures

Chapter 1

Connecting Alzheimer's disease, APOE-ϵ4 and resting-state networks: the role that MEG can play

Alzheimer's disease (AD) is the most prevalent form of dementia, accounting for 50-70% of all cases globally. It is characterised by a progressive loss of cognitive function [135]. Patients suffering from AD undergo a slow degeneration in their ability to care for themselves, eventually requiring full-time care in the later stages of the disease. This degeneration lays a tremendous emotional and financial burden on the patient and their family, as well as on the healthcare system.

1.1 The clinical significance of Alzheimer's disease

In 2006 there were an estimated 26.6 million people suffering from Alzheimer's disease globally [31]. The estimated global cost of all forms of dementia in 2003 was $156 billion for an estimated 27.7 million cases [137]. Over the next decades the prevalence of AD will increase dramatically as the world's population ages. Age is a major predictor of AD risk: AD has a prevalence of 1% in populations aged 65 and over, but 25% in populations aged over 80 years [54]. By 2050 it is predicted that there will be 106 million cases of AD globally (1 in 85 people) [31]. As of 2011, the National

Institute for Health and Clinical Excellence (NICE) recommends the combined use of three acetylcholinesterase inhibitors (*donepezil*, *galantamine* and *rivastigmine*) as pharmaceutical treatments for mild to moderate AD and the use of a NMDA receptor antagonist (*memantine*) for treating moderate to severe AD [100]. However, the benefits of all of these pharmaceutical interventions are limited at best [114]. If the average initial onset of Alzheimer's disease could be delayed by just one year, then this would lead to an estimated 9.2 million fewer cases in 2050 [31].

1.2 Current understanding of Alzheimer's disease

Numerous advances have been made in identifying key pathologies in AD. Three major factors have been found which are known to cause the neurological degeneration associated with AD: the amyloid precursor protein (APP), the microtubule-associated protein tau (MAPT), and the 4th isoform of apolipoprotein E (APOE4) [2]. Abnormal function of APP leads to its conversion to amyloid β peptide (Aβ) which in turn is known to form plaques. The hyperphosphorylation of MAPT causes neurofibrillary tangles (NFTs). APOE4 is a cholesterol transporting protein that has been proven to increase the risk of developing both late-onset and sporadic AD [41]. Better understanding of the APP, MAPT and APOE4 pathways to AD pathology would facilitate the development of biomarkers for early detection, which in turn would aid the design of new interventions to treat AD.

1.3 The importance of APOE-ϵ4

The APOE gene has three variants ϵ2, ϵ3 and ϵ4. Each individual has two versions of the APOE gene (which can be different variants or two copies of the same one). The overall frequencies of the APOE genes in western/Japanese populations (shown in Figure 1.1A) are ϵ2 - 5-10%, ϵ3 - 80% and ϵ4 - 10-15% [63]. The genetic predisposition towards development of Alzheimer's disease caused by APOE-ϵ4 was first characterised by Corder et al. in 1993 [41]. 95 patients affected by AD and 139 unaf-

fected subjects were studied to see the effect of APOE-ϵ4 on the risk of development of AD, age of onset and length of survival. It was found that a single copy of APOE-ϵ4 increased the risk of developing AD by 2.84 times, and two copies increased the risk by 8.07 times ($p < 0.00001$) compared with subjects without APOE-ϵ4 (see Figure 1.1B). In addition, the mean age of onset for subjects without the ϵ4 variant was 84.3 (1.3 standard error of the mean [SEM]); for subjects with one APOE-ϵ4 gene the mean age of onset was 75.5 (1.0 SEM) and finally for subjects with two copies of APOE-ϵ4 it was 68.4 (1.2 SEM). Similarly mean age of survival was 84.9, 78.8 and 78.1 for patients with 0, 1 and 2 copies of APOE-ϵ4. Corder et al. do not explain why the APOE-ϵ4 gene increases predisposition towards AD. However, they do note that 19 out of 95 AD patients did not have any copies of APOE-ϵ4, showing that it was not the sole factor in developing AD. In addition, patients with no APOE-ϵ4 came from families with AD predisposition suggesting that other genetic links other than APOE-ϵ4 affect the development of AD.

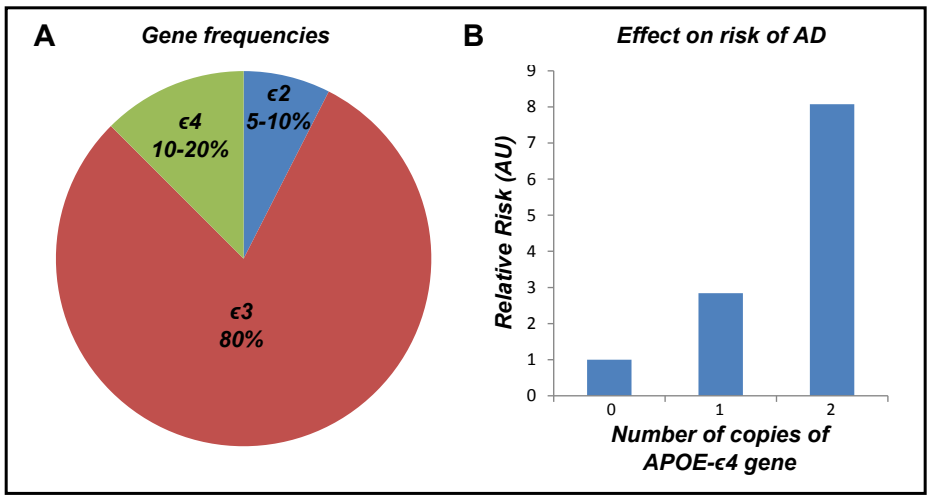


Figure 1.1: *Graphs showing A. The frequencies of occurrence of the different alleles of APOE and B. The effect that the genetic dose of APOE-ϵ4 has on one's risk of developing Alzheimer's disease. Values for A. and B. taken from [63] and [41] respectively.*

1.4 A comparison of neuroimaging modalities

In the previous sections, we outlined the clinical setting of AD. The pathology of AD is extremely complicated: consequently AD research involves efforts across a range

of disciplines including, but not limited to, biochemistry [69], genetics [41] and behavioural psychology [35]. Neuroimaging has a large role to play in understanding the onset and development of AD *in vivo*. In this section, we briefly compare some of the most commonly used neuroimaging modalities, contrasting their insights, strengths and limitations: every modality offers a trade-off between spatial resolution, temporal resolution, clinical applicability, invasiveness and cost. Figure 1.2 provides a concise summary of the trade-off between spatial resolution, temporal resolution and invasiveness offered by the main, currently available, human, *in vivo* neuroimaging techniques [70].

Functional Magnetic Resonance Imaging (fMRI): fMRI measures the changes in local magnetic permeability introduced by the deoxygenation of haemoglobin. Most fMRI studies measure the blood-oxygen-level-dependent (BOLD) signal, which reflects spatially localised changes in metabolic demand from neural populations. Among the non-invasive modalities, it has the best spatial resolution (of the order of 1mm), making it an excellent tool for spatially mapping activity across the brain. However, the BOLD signal is an indirect measure of neuronal activity that is a function of cerebral blood flow, cerebral blood volume and the metabolic rate of oxygen consumption [46]. These variables are collectively modelled by the haemodynamic response function (HRF). Spatial variability in all of the parameters making up the HRF cause it to vary greatly across the brain. A consequence of the HRF is that BOLD has a fundamentally poor temporal resolution of the order of seconds.

Positron Emission Tomography (PET): PET is an alternative modality for measuring the metabolic impact of neural activity by introducing a radioisotope which decays to produce positrons. Each collision of a positron with a surrounding electron (and subsequent annihilation) releases two high energy photons which can be detected outside the head. By detecting these annihilation events, a functional map of the uptake of the radioisotope can be estimated. PET was the modality used in the initial discovery of the default mode network [112]. It has poorer spatial resolution

than fMRI but can be used to estimate absolute quantitative measures of metabolism more readily. PET does require the administration of a radioisotope, which places constraints on how widely and regularly it can be used. However, by changing the radioisotope used, one can image a range of different biochemical processes such as glucose metabolism (using FDG-PET) or Aβ deposition (using PIB-PET) [32, 84].

Invasive Intracranial Recordings: There is a range of invasive techniques for recording neural activity (such as single unit recordings and electrocorticography (ECoG)). These techniques typically have excellent spatial precision and temporal resolution, but have limited coverage of the brain. Furthermore, because they are highly invasive they can only be applied to heterogeneous clinical populations (e.g. for localising sources of seizures in epilepsy patients) and usually in small sample sizes [70].

Electroencephalography (EEG): EEG measures the surface potentials due to volume currents generated by neural activity. It is a low cost, portable and highly robust neuroimaging tool with excellent temporal resolution. However, because the brain, skull and scalp have highly inhomogeneous conductivity profiles, it is extremely difficult to infer the spatial pattern of neural activity that drives an externally measured EEG recording. Consequently, most EEG studies are limited to sensor-level analyses, which limits their interpretability. That said, EEG has been successfully combined with fMRI via concurrent recordings of both modalities. By correlating the sensor-space band limited power of the EEG with the voxel-level BOLD time series, one can spatially localise the neural activity measured across multiple frequency bands [22, 91, 99].

Magnetoencephalography (MEG): MEG measures the magnetic fields outside the head that are generated by neural activity, with millisecond temporal resolution. It offers a complementary insight to EEG. However, in MEG we can localise the neuronal sources that drive our external recording with much more accuracy than in EEG. This gives MEG a superior spatial resolution (of the order of 5mm in cortical

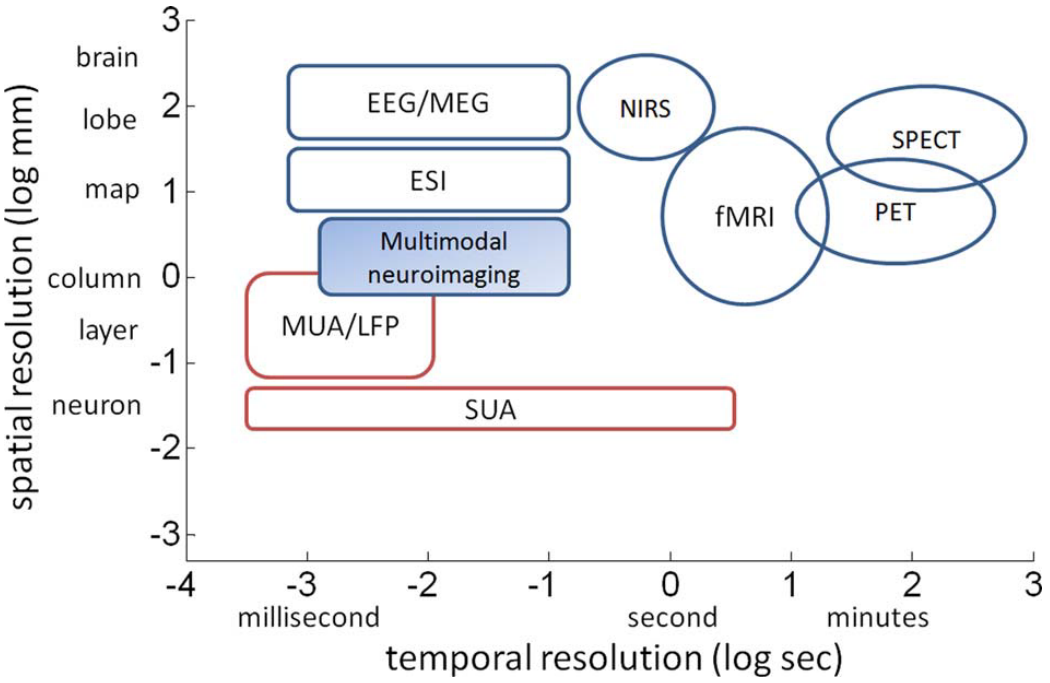


Figure 1.2: *Comparing neuroimaging modalities by temporal and spatial resolution. Invasive techniques have red borders. Electrophysiological imaging techniques are shown in oblongs. Non-neuronal imaging techniques are shown in ovals. SUA: single unit activity; MUA/LFP: multi-unit activity/local field potentials; ESI: electrophysiological source imaging corresponds to source-reconstructed analyses of EEG/MEG: electroencephalography/magnetoencephalography; NIRS: near-infrared spectroscopy; fMRI: functional magnetic resonance imaging; PET: positron emission tomography; SPECT: single-photon emission computed tomography. Figure taken from [70, Fig. 1].*

areas). MEG's spatial resolution is poorer than that of fMRI, but the signals it records are not confounded by the haemodynamic response function.

In summary, MEG offers a unique insight into resting-state spontaneous brain activity. Its high temporal resolution, reasonable spatial resolution and direct measurement of neural activity enable us to observe the mechanisms that drive cognitive function in a way that captures the rich spatio-temporal dynamics of brain function.

1.5 Functional connectivity and Alzheimer's disease

Within neuroimaging, there are a range of techniques for investigating Alzheimer's disease. Approaches include: structural studies (looking at grey matter atrophy [47]), changes in white matter connectivity (using diffusion-based imaging techniques [23]) or metabolic investigations (using various types of PET imaging [32]). In this thesis, we investigate the effect of the APOE-ϵ4 gene using concepts from the field of

functional connectivity.

1.5.1 What is functional connectivity?

Friston (2011) defined functional connectivity (FC) as "*statistical dependencies among remote neurophysiological events*" [59]. Healthy brain function requires the successful communication and integration of multiple spatially separated regions [132]. FC analysis is a statistical tool-kit for mapping these interactions in space and time. Within this framework, we can investigate how external stimuli (such as task-positive activity or disease) mediate these functional relationships. FC can be used to measure *direct* and *indirect* interactions between regions: *direct* FC is a statistical dependency resulting from direct interaction between two regions; *indirect* FC corresponds to the statistical relationship between two regions which do not communicate directly but which have a statistical co-dependency due to interactions mediated through a third (or more) region(s) [124].

Functional connectivity should not be confused with *effective connectivity* which attempts to elucidate the causal nature of the statistical relationship observed in FC. Typically effective connectivity is estimated using more mechanistic, biophysically motivated approaches, such as dynamic causal modelling and structural equation modelling [58, 94]. True measures of effective connectivity would provide a much more informative picture of human brain function and dysfunction. Unfortunately, measuring resting-state effective connectivity is an extremely challenging task, particularly when considering large numbers of nodes ($>$10) as we do in this thesis, and, as such, is beyond the scope of the work presented here.

1.5.2 Functional connectivity and the resting-state

We can measure functional connectivity under a range of conditions but will focus primarily on functional connectivity during the resting-state (although chapter 6 develops a framework for analysing task-positive MEG data). The resting-state paradigm involves scanning a subject in the absence of any goal-directed task or external stim-

ulation [112]. Typically, the subject is either scanned with their eyes closed or with their eyes open but fixating on a cross. This simple paradigm has been extremely useful for the following reasons: it translates extremely well across imaging modalities, healthy control and patient populations, and has very few design parameters. Despite its simplicity, the resting-state can be highly useful for measuring functional connectivity. Biswal et al. measured resting-state functional connectivity in fMRI by measuring the correlation between the BOLD signal of a seed region (identified with a motor localiser task) and the rest of the brain [21]. This analysis identified a group of several regions, all associated with motor processing. This was the first example of imaging a resting-state network (RSN) in fMRI.

RSNs are spatially separated regions of the brain which have temporally covarying neural activity indicative of some functional integration. The identification of the sensorimotor network (SMN) by Biswal et al. represented the first step in a process to describe resting-state BOLD activity as the aggregated activity of multiple networks. In 2001, Raichle et al. used PET to measure the oxygen ejection fraction of resting-state brain activity [112]. They found a set of regions that shared the property of being consistently anti-correlated with task activity. This network was termed the default mode network (DMN). In 2005, Fox et al. demonstrated that the default mode network could be extracted by estimating the correlation in the BOLD signal between a seed region in the posterior cingulate cortex (PCC) and all other voxels [56]. They also found the dorsal attention network (DAN) which was found to be consistently anti-correlated with the DMN.

Up until this point, seed-based correlation was the standard technique for imaging these networks (seed-based correlation methods in MEG are discussed in section 4.4.2). However, seed-based correlation is limited to imaging single networks at a time and is extremely sensitive to the choice of seed voxel. In 2005, Beckmann et al. applied independent component analysis (ICA) to extract multiple RSNs from resting-state fMRI time series [13]. ICA is a blind source separation technique that decomposes a matrix (that is a mixture of signals) into a set of underlying, statistically independent sources and a mixing matrix [16,40,97]. A more thorough introduction to ICA is given

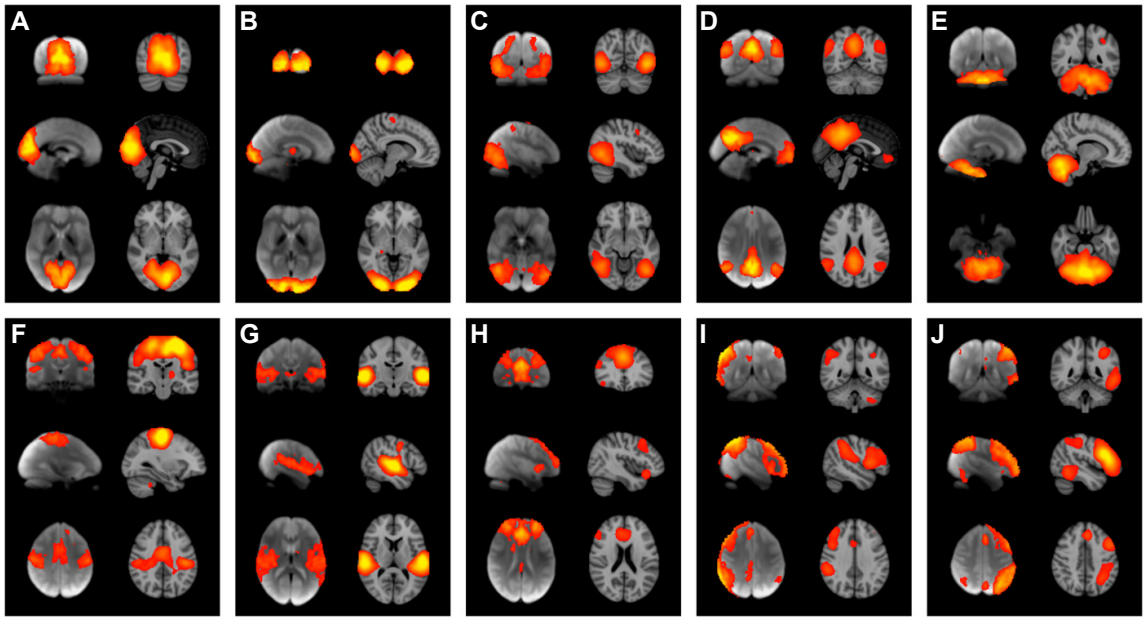


Figure 1.3: *Figure of 10 RSNs imaged using a spatial ICA applied to temporally concatenated BOLD time series from 36 subjects. For each panel: Z-statistical images (thresholded at $Z = 3$) of the RSN maps are on the left. Corresponding task-activation maps from the BrainMap database are on the right. RSNs correspond to A. Medial Visual Network; B. Visual Network; C. Lateral Visual Network; D. Default Mode Network; E. Cerebellum; F. Sensorimotor Network; G. Auditory Network; H. Executive Control Network; I. Right Frontoparietal Network; J. Left Frontoparietal Network. Figure taken from [122, Fig. 1].*

in section 4.5. Beckmann et al. used ICA to decompose the temporally concatenated BOLD time courses from a cohort of 10 subjects into a set of spatially independent maps corresponding to multiple known RSNs. In 2009, Smith et al. performed a similar analysis on 36 subjects to obtain 20 spatially independent networks [122]. They selected 10 RSNs from the ICA output and demonstrated, via a meta-analytic comparison with the BrainMap database of task activations, that these RSNs were spatially associated with regions that are known to be involved in behavioural function. Figure 1.3 shows the 10 RSNs and their corresponding task activation maps. Although RSNs have been studied in multiple modalities, their characterisation in fMRI represents the best available consensus definition. Consequently, Figure 1.3 is used as the benchmark definition for RSNs throughout this thesis.[1]

[1]The author would like to thank Professor Steve Smith (FMRIB, Oxford) for providing the original 20 spatial maps which are used throughout this thesis.

1.5.3 Electrophysiological characterisations of RSNs

In recent years, several studies have investigated resting-state functional connectivity using electrophysiological neuroimaging techniques, including concurrent EEG/fMRI and MEG. It is critical to note at this point that all sensor and source-space EEG and MEG FC analyses are contaminated with spurious contributions to FC. In sensor-space, this is referred to as volume conduction (for EEG) or field spread (for MEG). In source-space, this is known as cross-talk or signal leakage between voxels. We will discuss these issues in more detail in sections 2.6 and 4.4.1. However, when considering any FC analysis in EEG or MEG, accounting for these spurious contributions is essential.

Mantini et al. (2007) [91] used concurrent EEG/fMRI to show that the electrophysiological recordings found in EEG correlate with the component time courses found in fMRI spatial ICA. In addition, they showed that this correlation is found in low frequency fluctuations in the power of the EEG signals across all bands. They showed that the default mode network (positively) correlates most strongly with the alpha and beta band EEG signals. However, they postulated that all frequency bands should be considered together in order to properly capture the relationship between the resting-state BOLD and EEG signals.

In 2010, de Pasquale et al. used MEG to find the dorsal attention network (DAN) and the default mode network (DMN) [43]. Initially using standard seed-based correlation, a comparison was made between the DMN and DAN using fMRI and MEG. Networks, found with a single seed and using wideband (1-150Hz) MEG, were poorly defined and unrealistically confined to the ipsilateral hemisphere and predominantly around the seed region. This result assumed temporal stationarity in the MEG data. To improve the detection of RSNs in MEG, they used the *maximally-correlated window approach (MCW)* that compensates for non-stationarity in FC. The resting data were epoched into regions of high correlation at three principle nodes of each network (referenced against a fourth unconnected node). This method enabled the seed-based correlation to detect the DAN/DMN in both hemispheres. De Pasquale et al. found

that the best spatial localisation was achieved in the alpha, beta and theta frequency bands compared with the gamma band. In this approach, there was no method for accounting for signal leakage between voxels, which is expected to produce artificial correlation especially near the seed region.

Liu et al. (2010) analysed a range of different resting states at the sensor level (including eyes-open, eyes-closed and light sleep) using MEG [86]. They found that very low frequency modulations in the power of all the frequency bands could be detected (by estimating the amplitude envelope via the Hilbert transform). They also found that inter-hemispheric correlations peaked in the low beta band. They did not explicitly account for effects of field spread and so their finding may have been partially confounded by spurious correlations.

Brookes et al. (2011) measured inter-hemispheric functional connectivity within the sensorimotor network (SMN) using MEG [25]. Six subjects were scanned using fMRI during a finger tapping exercise to determine the seed position in the left motor cortex. The same six subjects were scanned with MEG and a beamformer used to reconstruct their source-space time courses across the whole brain over 8 frequency bands. The envelope time course at every voxel was estimated via the Hilbert transform and then temporally down-sampled using a non-overlapping windowed average. Optimal correlation between left and right motor areas was achieved in the low beta band (13-20Hz). Longer down-sampling windows were shown to boost the inter-hemispheric connectivity between the left and right motor cortices. Spurious contributions to FC due to signal leakage were accounted for through Monte Carlo simulations of uncorrelated data, which allowed a probabilistic threshold to be determined that excluded likely artefactual correlations.

Brookes et al. followed up on this work by performing a temporal ICA decomposition of the group concatenated envelopes of source-reconstructed MEG data to estimate MEG-derived RSNs [30]. They were able to find 8 MEG analogues of the standard fMRI RSNs (shown in Figure 1.4): 7 from the beta band and 1 from the alpha band (specifically the DMN). This study represented the first fully independent electrophysiological imaging of multiple RSNs.

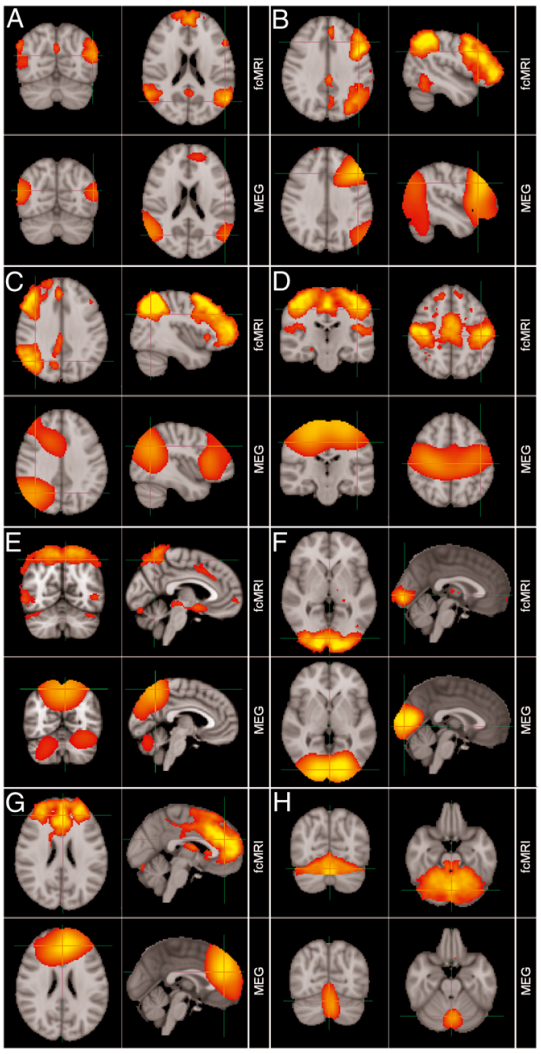


Figure 1.4: *Figure of 8 RSNs imaged by applying temporal ICA to the temporally concatenated source-space envelopes from 10 subjects. For each panel: fMRI spatial ICA derived Z-statistical images (thresholded at $Z = 3$) of the RSN maps are shown above [122]. Corresponding MEG RSN correlation maps are shown below. A. Default Mode Network (α-band); B. Left Frontoparietal Network; C. Right Frontoparietal Network; D. Sensorimotor Network; E. Superior Parietal Lobule; F. Visual Network; G. Frontal Network; H. Cerebellum. Maps B-H were estimated from the envelopes of the β-band. Figure taken from [30, Fig. 1].*

More recently, MEG studies have started to move beyond validating fMRI RSNs using a direct electrophysiological measure. De Pasquale et al. (2012) used their maximally-correlated window technique to find evidence that, when the DMN (in particular the posterior cingulate cortex) is highly engaged with itself, it also demonstrates strong cross-network interactions [44]. Hipp et al. (2012) demonstrated that graph theoretical measures such as *betweenness* and *degree*, which capture the network structure of whole-brain FC, exhibit strong spatial and frequency specificity [74].

1.5.4 Functional connectivity, APOE-ϵ4 and the default mode network

The pathology of Alzheimer's disease is extremely diverse and includes grey matter atrophy [47], abnormal metabolism [96], and Aβ plaque deposition [69]. Furthermore, AD-related pathology is not localised to a single region but instead is found in multiple brain areas. As such, developing a causal hypothesis that links APOE-ϵ4 to an increased risk of AD is not trivial. This is especially true when considering AD patients whose brains have undergone such extensive AD-related damage. To investigate the causal role that APOE-ϵ4 may have on healthy brain function resulting in a predisposition to developing AD, several studies have focused on its effect on healthy populations who are well beneath the age of typical AD onset [55, 129].

An increasingly important concept in understanding the development of AD is the relationship been the default mode network and AD pathology. Greicius et al. (2004) showed that the DMN in AD patients was different to healthy ageing adults [64]. Using ICA to extract the DMN from resting-state fMRI scans of the two groups, they compared the goodness-of-fit between each group's DMN and a reference DMN map from an independent data set. They found significantly poorer goodness-of-fit for the AD population than for the healthy ageing group, suggesting that abnormal DMN functional connectivity is associated with AD. Although interesting, this study did not provide evidence of abnormal FC having a causal influence on AD. Buckner et al. (2005) showed an overlap between Aβ deposition (measured with PIB-PET), abnormal metabolism (measured with FDG-PET), and grey matter atrophy (measured with MRI) in AD patients and the DMN measured in a young healthy control group. This further implicated the DMN in AD pathology but again could not establish whether the DMN dysfunction causes AD pathology or if the correlations are simply epiphenomenal [32].

Both Greicius et al.'s fMRI study and Buckner et al.'s multi-modal study build toward a hypothesis that implicates abnormal DMN activity with AD. Abnormal metabolism in the posterior cingulate cortex is one of the most consistent obser-

vations in PET imaging of early-stage AD [96]. However, the major behavioural and pathological changes in AD are associated with medial temporal lobe damage. Therefore, abnormal DMN function is an excellent candidate for linking the observed differences in the two regions.

To investigate a causal link between abnormal DMN function and AD, several studies have looked at whether the APOE-ϵ4 gene affects functionally-connected networks, such as the DMN [55, 129]. Using ICA and dual regression to look at RSN-associated group differences, Filippini et al. (2009) showed abnormal functional connectivity in the hippocampal nodes of the DMN in a healthy group of APOE-ϵ4 carriers who were beneath the typical age of AD onset. Trachtenberg et al. (2012) found additional differences in APOE-ϵ4 carriers, associated with the auditory network and left frontoparietal network. These studies show that abnormal DMN activity is present in groups with a greater risk for AD. However, to fully demonstrate a causal link between abnormal RSN functional connectivity would require a very long term longitudinal study.

All of the analyses investigating the relationship between abnormal functional connectivity and AD pathology have been based on fMRI and PET [32, 55, 64, 129]. Both these modalities provide very indirect measures of brain activity. The fMRI blood-oxygen-level-dependent (BOLD) signal and the PET signal are both intrinsically confounded by vascular processes. Therefore it is impossible to determine whether any relationships observed between AD, APOE-ϵ4 and the default mode network are due to neuronal or vascular processes. In addition, fMRI and PET have a relatively poor temporal resolution of the order of seconds, limiting any studies to looking at low frequency activity. MEG has the advantage of directly probing neural activity at the millisecond time scale, allowing a much wider range of temporal dynamics to be explored.

1.6 Summary

The role of APOE-ϵ4 on brain function and its contribution to AD is an ongoing question. Functionally-connected networks are a strong candidate for linking APOE-ϵ4 to ultimate AD pathology. MEG allows us to investigate the electrophysiological basis of the effect that APOE-ϵ4 exerts on healthy brain FC. That said, MEG FC analysis has some unique and complex challenges that have to be addressed. This thesis presents a set of techniques for measuring functional connectivity in MEG including the following key novel contributions to the field of MEG FC analysis. Chapter 2 discusses in more detail what MEG actually measures and how we can use it to measure underlying neural activity. In chapters 3 and 4 we present a framework for preprocessing resting-state MEG data and imaging RSNs. We have made critical progress in developing tools for MEG analysis that are compatible with data acquired on *Elekta Neuromag* systems. This includes techniques for combining multiple sensor types and accounting for signal space separation during both ICA de-noising and beamforming. We present the first successful imaging of RSNs from *Elekta Neuromag* data. In chapter 5, we present a framework for detecting differences in resting-state networks between populations and apply it to a data set of healthy people who have been grouped by their APOE genotype. We show that APOE-ϵ4 carriers have significant differences in oscillatory power in both the left and right frontoparietal networks. In this chapter we present a critical analysis that shows how beamformer *weights normalisation* introduces a major confound into variance-based statistics across multiple sessions of separately beamformed data. Chapter 6 describes our approach for investigating functionally-connected networks in task-positive data. We show that our method, which couples ICA with the General Linear Model, can provide useful insights into task-related functional networks. Chapter 7 introduces our preliminary investigation into parcellation and partial correlation network analysis techniques for MEG. As part of this, we develop a novel method for accounting for signal leakage between multiple parcels. In chapter 8, we conclude this thesis by considering future avenues of research that could build on the work presented here.

Chapter 2

An introduction to MEG

In the previous chapter, we have outlined the unique insight that magnetoencephalography (MEG) can bring when investigating interactions between the APOE-ϵ4 gene and functional brain connectivity in the resting-state. In this chapter, we consider the physiology of the underlying neural activity that creates the MEG signal, provide an overview of the instrumentation used to detect it, and introduce the techniques that are used to reconstruct the neural activity.

2.1 Origins of the magnetoencephalogram

MEG measures direct neuronal activity. However, to understand what this means in more detail it is necessary to understand how the electrophysiological behaviour of an individual neuron interacts with other neurons to form large and coherent neuronal assemblies, which are capable of generating neural oscillations of sufficient amplitude to produce a detectable magnetoencephalogram.

We must first define what we mean by detectable. Hämäläinen et al. stated that a current dipole in excess of 10nAm is necessary to generate a detectable magnetic field [66]. In the following sections we will show that this requires the synchronous activity of tens of thousands of strongly aligned neurons. By building up a model from the single neuron level to larger neural assemblies, we can show that human brain function falls within these constraints.

2.1.1 Electrophysiological function at the single neuron scale

Neurons propagate information via a combination of electrical and chemical processes. The cell membrane of a neuron contains multiple ion channels that, when open, allow charged ions to cross the cell membrane. These channels can either be *voltage-gated*, which means that they open when the local membrane potential exceeds a threshold, or *ligand-gated*, which means that they open when a specific neurotransmitter binds to them. If an ion channel opens, it depolarises the surrounding patch of cell membrane which in turn causes any neighbouring voltage-gated ion channels to open, depolarising the next patch of membrane. It is by this mechanism, known as an *action potential*, that neurons send information along their basal dendrites to other neurons. Where one neuron terminates and another begins is known as a *synapse*. When an action potential arrives at a synapse, it triggers the release of neurotransmitters that bind with ligand-gated channels in the next neuron. It is by this mechanism that signals are passed between neurons.

To conceptualise how these mechanisms translate into a measurable magnetoencephalogram, we start by considering a single neuron in isolation, specifically focusing on pyramidal neurons. Pyramidal neurons have a generic structure outlined in Figure 2.1B, with four main parts:

- **Soma**: The cell body containing the nucleus.
- **Apical Dendrite**: This cell structure contains all the synapses with incoming neurons.
- **Axonal Hillock**: This structure initiates an action potential when its local potential exceeds its firing threshold.
- **Basal Dendrite**: This long and branching structure transmits resulting action potentials to synapses with other neurons.

For a given neuron to generate an action potential, the cell membrane potential at the axonal hillock must rise from its resting potential to exceed a threshold potential. This causes voltage-gated ion channels in the cell membrane to open, initiating an

action potential that propagates down the cell's axon to terminate at synapses with other neurons. The probability of a neuron generating an action potential is mediated by the synapses between other neurons and the apical dendrite. Excitatory synapses induce excitatory post-synaptic potentials (EPSPs) which raise the overall membrane potential, increasing the likelihood of an action potential, whilst inhibitory synapses induce inhibitory post-synaptic potentials (IPSPs) which lower the membrane potential and decrease the likelihood of an action potential occurring. In more detail, consider the following sequence of events (shown in Figure 2.1B):

1. An action potential arrives at the axon terminal of an excitatory or inhibitory synapse between the terminating neuron and the apical dendrite of the receiving neuron.

2. Neurotransmitters (green circles) are released from the axon terminal and bind with ligand-gated ion channels in the apical dendrite's synaptic membrane.

3. Ligand-gated ion channels open to allow the flow of positive ions (red crosses) **into** the cell for an excitatory synapse/**out of** the cell for an inhibitory synapse.

4. Excitatory: Flow of positive ions into the cell causes an increase in the cell membrane potential in the region of the synapse. Inhibitory: Flow of positive ions out of the cell causes an decrease in the cell membrane potential in the region of the synapse.

The cumulative effect of all the inhibitory and excitatory synapses modulates the neuron's ability to initiate an action potential. However, in terms of understanding the generation of an MEG signal, we are not concerned with the neuron's generation of an action potential, but with the resulting intracellular currents generated by the EPSPs and IPSPs, which flow along to the apical dendrite (in opposing directions) [68]. Estimating the exact current dipole generated by a single post-synaptic potential is not trivial but a combination of *in vitro* experiments and biophysical simulations estimate it to be in the range of 0.02-0.90pAm [66, 68].

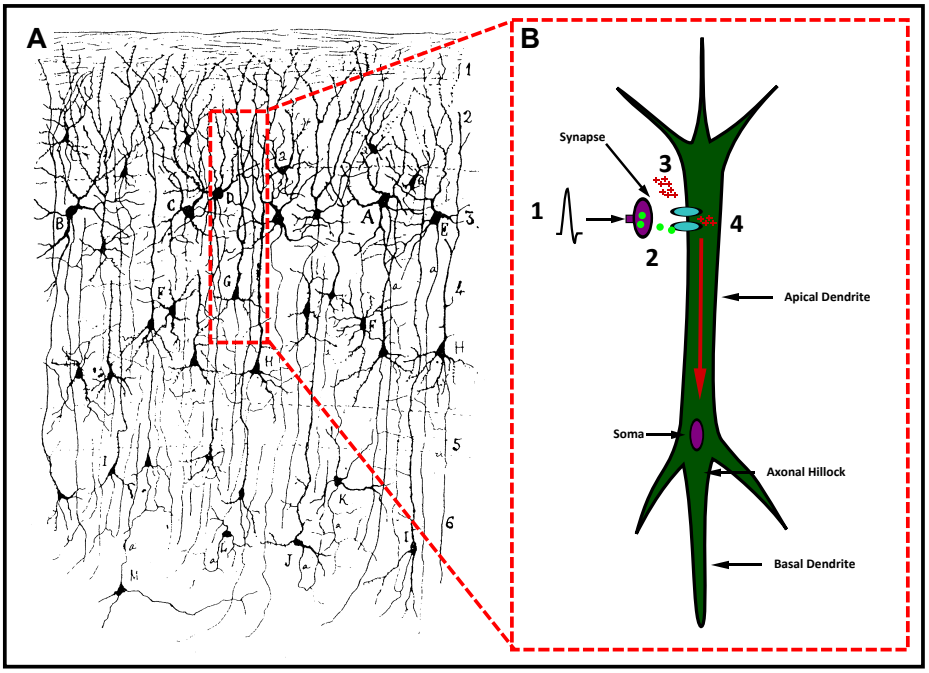


Figure 2.1: *A. Histological cross section of the cortex of a rat, showing the high degree of alignment of pyramidal neurons. B. A schematic of the process by which post-synaptic potentials are generated in pyramidal neurons.* ***1****. Incoming action potential causes* ***2****. the release of neurotransmitters that bind to ligand-gated membrane channels which open and allow* ***3****. the flow of positive ions across the cell membrane leading to* ***4****. intracellular currents along the apical dendrite. Figure A is taken from [66, Fig. 14].*

2.1.2 Generation of rhythmic brain activity

In the previous section, we demonstrated a regime in which a single neuron behaved like (and therefore can be modelled as) a current dipole. However, evidence (*in vivo* and *in silico*) indicates that current dipoles resulting from **PSP**s are much too small to be detected individually [68]. An ensemble of the order of 50,000 current dipoles due to **PSP**s would be needed to cumulatively produce a current dipole moment that could be detected with an MEG system. Furthermore, the signals recorded from a resting-state MEG scan are strongly oscillatory, reflecting the same neural oscillations seen in EEG, such as the dominant alpha band oscillations (8-13Hz). Cortico-thalamic networks have been suggested as a mechanism by which large scale, synchronised oscillatory behaviour can emerge from high frequency, independent neuronal activity [33,116]. This resonant behaviour recruits sufficient numbers of synchronised neurons to generate a measurable oscillatory MEG signal.

A critical question is whether these oscillations are simply some epiphenomenal

process that arises as consequence of some non-oscillatory neuronal behaviour that is responsible for cognitive function. Several different theories have emerged suggesting that neural oscillations do actually play a key role in cognition.

Considering the alpha oscillation (the dominant oscillatory frequency band observed in the resting-state), there are several hypotheses. Pfurtscheller et al. (1996) suggested that alpha oscillations represents some process of cortical idling [110]. An alternative theory characterises the alpha rhythm as the brain's pacemaker [33]. More recently, the alpha rhythm has been suggested to play a role in *gating by inhibition*: the process by which the brain shuts off task-irrelevant pathways in response to external stimuli [82].

Of course, the alpha band is not the only frequency band to demonstrate oscillatory synchronisation. Other oscillatory bands, such as the theta band (4-8Hz) and gamma band (>30Hz) have been implicated in facilitating neural communication at different spatial and temporal scales in the resting-state and in task-directed activity [33].

2.1.3 Structure and arrangement of the human cortex

The dipoles induced by a single cell's activity are insufficient to produce magnetic fields that can be detected externally. In order to produce a detectable field, a group of neighbouring neurons must generate synchronised current dipoles which sum to a single current dipole in excess of 10nAm [66]. However, synchronicity is insufficient as two aligned but opposed dipoles of equal magnitude will cancel each other out. More formally, the current dipoles of individual neurons must collectively form an open field [68]. Fortunately, the natural cyto-architecture of the human brain produces the required open fields. The cortex of the human brain is effectively a sheet of pyramidal neurons with strongly aligned apical dendrites that has been folded to fit inside the skull. This folding is remarkably consistent between individuals and provides sufficient structural alignment between the apical dendrites of pyramidal neurons to allow for normal brain function to generate a detectable magnetoencephalogram. Figure 2.1A shows a slice through the rat cortex showing the strong degree of structural alignment

in the mammalian cortex [66, Fig. 14].

However, although the human cortex shows extremely uniform neuronal alignment, MEG has a strong sensitivity to specific regions of the cortex. Sarvas showed that a current dipole inside an homogeneous spherical conductor only produces an external magnetic field if it has a tangential component [117]. This is because the return currents of the radial component cancel out the field generated by the current dipole. Though the brain is not an homogeneous sphere, it is sufficiently similar that the MEG sensors will be predominantly sensitive to the tangential component of a current dipole and insensitive to its radial component. When considering the folded sheet that is the human cortex, sources in the gyri and sulci, which are mostly radial, will produce weak external magnetic fields whilst sources located between the gyri and sulci will produce stronger fields because they have a large tangential component [66, 68].

2.1.4 Deep brain sources

Until this point, it has been assumed that all the neuronal activity recorded in MEG originates from the cortex. Some source reconstruction techniques utilise this feature and constrain all reconstructed neural activity to the cortical surface. However, many deep brain structures (such as the hippocampus) are of great interest and it is critical to understand if and how deep brain sources can be detected. Deep brain regions like the hippocampus do not exhibit optimal neuronal alignment of apical dendrites of pyramidal neurons, which leads to less open fields. Furthermore, because of their depth, they are further from the MEG sensors and so have a substantially weaker measured signal (the drop off of magnetic flux as a function of distance from the dipole source follows the inverse square law).

Considering the case of the hippocampus, a region that is of great interest in AD, Quraan et al. (2001) developed simulations suggesting that under suitable conditions, including the use of a suitable experimental paradigm, the correct source reconstruction techniques and sufficient trials, hippocampal activity can be localised with MEG [111]. This finding has been empirically supported by other findings and

suggests that whilst cortical sources will be the dominant contributor to the magnetoencephalogram, non-cortical neural activity is potentially detectable [88].

2.2 Detecting the MEG signal

In order to detect the magnetic fields generated by human brain function, a sensor capable of resolving signals of the order of 10 femtoteslas is required. The first attempt to record the human magnetoencephalogram was by Cohen et al. in 1968. They successfully detected the human alpha rhythm with a single 1 million turn coil with a ferrite core [37]. However, the application of super conducting quantum interference devices (**SQUID**s), by Cohen in 1972, represented the major technological leap in MEG [38]. **SQUID**s have the following generic structure: a loop of superconducting material broken by a very thin, non-conducting layer (known as a Josephson junction). The superconducting loop by itself behaves in a perfectly diamagnetic fashion. Any magnetic field passing through the loop induces a current inside the loop. The induced current generates a magnetic field in the material that opposes and cancels out the external field. However, measuring this current is not possible. The Josephson junction does not prevent the flow of induced currents but when a constant bias current is placed across the loop, the Josephson junction creates a flux dependent resistance which can be measured via the voltage across the loop [68].

Since Cohen's initial work, whole-head MEG systems have been developed that use arrays of the order of 300 **SQUID**s to sample the magnetic brain activity. The **SQUID**s do not detect the brain activity directly. Instead, for each **SQUID**, a flux transformer is used to capture magnetic flux and pass it to the **SQUID**. A flux transformer consists of a single superconducting wire formed of a pick-up coil, a compensation coil and a signal coil. The pick-up coil and compensation coil capture the external flux, which induces an internal current. This current generates an associated magnetic field at the signal coil which is in turn measured by the **SQUID**.

By using different flux transformer geometries, different types of MEG sensors can be designed with associated spatial sensitivity and interference rejection properties.

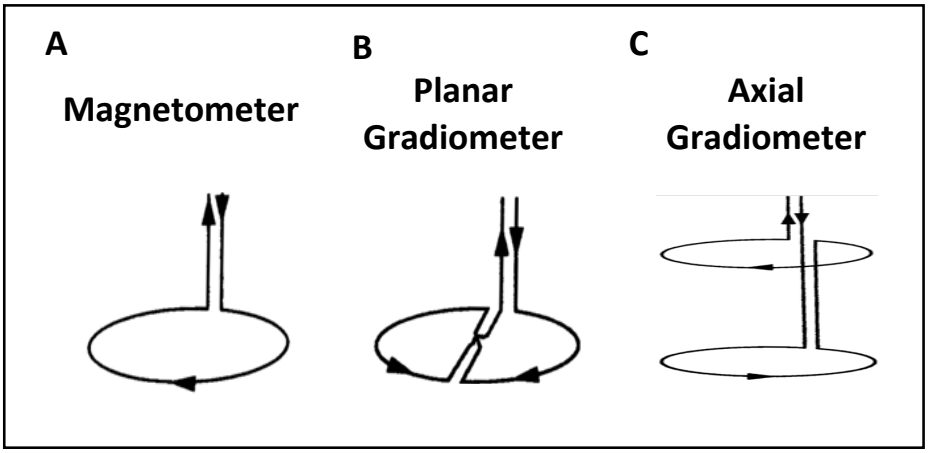


Figure 2.2: *Diagrams of the pick-up coils for the three main sensor types A. Magnetometer; B. Planar Gradiometer; C. Axial Gradiometer. Diagrams of the pick-up coils taken from [68, Fig. 2-5].*

Specifically three sensor types are considered here: 1.) **Magnetometers**, 2.) **Planar Gradiometers**, 3.) **Axial Gradiometers**. Fundamentally, these three configurations use the same physics of **SQUID**s. However, the variations in their design lead to notably different spatially sensitivity profiles and interference rejection characteristics. Magnetometers measure the normal component of the magnetic field passing through the pick-up coils. They have poorer spatial specificity but superior sensitivity when measuring deep brain sources. Planar gradiometers measure tangential components of the gradient of the normal magnetic field. They allow for improved localisation but are limited to imaging regions near the surface of the head (i.e. the cortex). Axial gradiometers measure the radial gradient of the normal magnetic field and have a slightly increased sensitivity to deeper sources. Figures 2.2A,B,C show how these three sensor types are configured. Gradiometers measure the difference between the flux passing through the pick-up and compensation coils, whereas magnetometers have no compensation coil and so directly measure the flux though the pick-up coil. Gradiometer design renders the sensors insensitive to far-field sources, which is ideal as far-field sources are caused by non-neuronal interference. However, their spatial selectivity also makes the gradiometer class of sensors less sensitive to deep brain sources.

2.3 Environmental and physiological sources of interference

The human magnetoencephalogram, whilst detectable, has substantially weaker power when compared with most common sources of interference. For example the MEG signal is one billionth that of the Earth's static field. Figure 2.3 demonstrates how poor the signal-to-noise ratio of the MEG signal is compared with a range of common sources of interference. Magnetic shielding provides the primary line of interference rejection and helps suppress all environmental sources of noise and interference (such as electrical noise from laboratory equipment). However, the magnetically shielded room (MSR) cannot attenuate interference from sources inside the MSR. As such, rigorous scanning protocols are essential to ensure that the participant has no magnetic materials (including clothing, jewellery, make-up and dental implants). Strict adherence to suitable pre-scan screening can remove these sources of noise. Unfortunately, the final group of artefacts are physiologically generated magnetic signals which include the magnetocardiogram and magnetomyograms due to muscle activity and eye-blinks. These artefacts cannot be avoided and will be recorded during data acquisition. As Figure 2.3 shows, these signals are many times larger than the MEG signal. To remove physiological interference, signal processing techniques, from averaging, to filtering, to blind source separation, must be used. This will be discussed in more detail in chapter 3, where the data preprocessing pipeline is outlined, including the novel application of independent component analysis for de-noising multiple sensor types simultaneously.

2.4 The forward problem, inverse problem and source reconstruction in MEG

MEG data, in its acquired form, consists of a vector of data samples for every sensor. Each sensor's time series consists of a mixture of the measured magnetic flux coming from each neuronal dipole. The flux observed at each sensor due to a given dipole

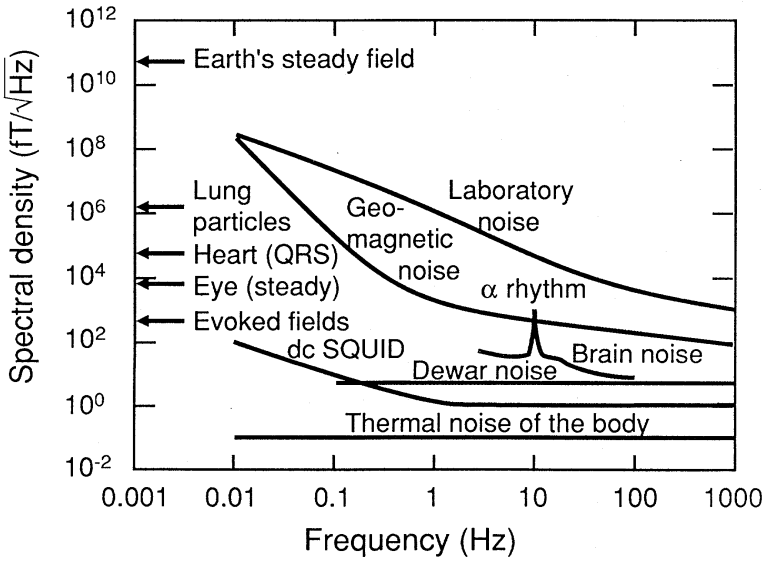


Figure 2.3: *Spectral density of the human magnetoencephalogram compared with common sources of environmental and physiological interference. Figure taken from [66, Fig. 8].*

is described by the quasi-static approximation of Maxwell's equations: the result of this approximation is that the flux induced by a current dipole of fixed orientation at a given time can be described as being proportional to the current of the dipole [66]. At a given time, t, a dipole at location q can be described with a 3 by 1 vector $\mathbf{x}_q(t) = [x_{q,i}(t), x_{q,j}(t), x_{q,k}(t)]^T$, where each element describes the net current flowing in the i, j and k directions. The flux, $y_{n,q}(t)$, observed at the n^{th} sensor due to this dipole is then described as the product of $\mathbf{x}_q(t)$ and $\mathbf{h}_{n,q}$. The lead-field $\mathbf{h}_{n,q} = [h_{n,q,i}, h_{n,q,j}, h_{n,q,k}]$ expresses the flux observed at the n^{th} sensor due to unit dipoles in the $[i, j, k]$ directions at location q. This is described mathematically by equation 2.1:

$$y_{n,q}(t) = \mathbf{h}_{n,q}\mathbf{x}_q(t) \tag{2.1}$$

Furthermore, the total flux observed at the sensors is the linear superposition of the flux induced by each current dipole. This description of the sensor-level data as a linear superposition of the underlying current sources is known as the MEG *forward*

problem and is mathematically expressed in equation 2.2:

$$\begin{aligned} y_n(t) &= \sum_{q=1}^{Q} \mathbf{h}_{n,q}\mathbf{x}_q(t) + \varepsilon(t) \\ \mathbf{Y} &= \sum_{q=1}^{Q} \mathbf{H}_q\mathbf{x}_q + \boldsymbol{\varepsilon} \end{aligned} \tag{2.2}$$

where $\mathbf{Y}$ and $\boldsymbol{\varepsilon}$ are the N sensors by T samples matrix of measured MEG data and uncorrelated sensor noise respectively. The N by 3 matrix $\mathbf{H}_q$ describes the MEG lead-fields for the q^{th} location in the brain.

To solve the *forward problem*, one needs only know the geometry of the head and the material properties in order to estimate the lead-fields $\mathbf{H}$. However, in MEG analysis, we are presented with the challenge of inferring the underlying brain activity from a set of external measurements, referred to as the MEG *inverse problem*. Solving the inverse problem is not trivial. MEG systems have around 300 sensors, while the brain can be modelled as a grid of the order of 10,000 dipoles, making the inverse problem highly ill-posed.

In the following sections, techniques for estimating the MEG lead-fields and approximating a solution to the inverse problem are described. This is collectively termed *MEG source reconstruction* and is a fundamental stage in the analysis of resting-state MEG data.[1]

2.4.1 Estimating MEG lead-fields

The lead-fields for a location are estimated by applying the quasi-static approximation of Maxwell's equations to the specific physical geometry of a given participant in a given scanner. Unfortunately exact solutions do not exist for this set of cases. In principle, numerical methods such as boundary element models, finite element

[1]It should be noted that sensor-level MEG analysis is an alternative approach which does not require source reconstruction techniques. Sensor-level analyses are particularly popular in task-positive paradigms due to their relative simplicity and for their equivalence to findings in EEG. However, this thesis is centred on the detection of functional connectivity which is generally best estimated after source reconstruction [118].

models and finite difference models could be applied (as have been used in EEG source reconstruction to estimate the voltage measured at the scalp due to a current source in the head). In EEG, numerical analyses are necessary as the currents that flow in the head follow highly complex paths depending on the local conductivity. These techniques are typically complex and computationally intensive. However, when considering magnetic fields, the permeability of the subject's head and the gap between the head and the pick-up coil are homogeneous and can be treated as equal to the permeability of free space. When combined with certain simplifying assumptions about the geometry of the head, analytical estimates for the lead-fields can be found. Three different methods are outlined here:

- **Single Sphere Model**: The single sphere model was developed by Sarvas [117], who considered a single, spherically symmetric volume containing a current dipole. Sarvas developed a solution for the magnetic field observed at a given location outside the volume due to the flow of currents within the sphere. This model yielded two key findings: firstly, that the radial component of the current dipole does not contribute to the field outside the sphere; secondly, that the field is independent of the (spherically symmetric) conductivity profile. By assuming that the permeability of the sphere is that of free space, Sarvas's model only requires the specification of the sphere geometry. This allows sufficiently accurate modelling of the flow of currents due to the neuronal dipole that give rise to the external magnetic field. This is achieved by fitting a sphere to the subject's head shape. The single sphere model is extremely useful for understanding the distribution of magnetic fields outside a body. However, modelling the head as a sphere is a very crude assumption and will limit the accuracy of the estimated lead fields.

- **Local Spheres Model**: Huang et al. (1999) proposed a more complex forward model approximation that models the head as multiple overlapping spheres [75]. For each sensor, an homogeneous sphere is fit to the head surface local to that sensor. The lead fields are estimated from the combination of all of these

spheres. This model has one important free parameter that defines the "local" patch of head surface for each sensor. Huang et al. compared the single sphere model and the local spheres model to a three shell boundary element model (BEM). The single sphere model fits poorly over the frontal and visual regions (intuitive as these regions are where the head shape diverges most strongly from spherical), whereas the local spheres model has much better agreement with the 3-layer BEM but does not require the computational expense.

- **Single Shell Model**: Nolte et al. proposed using a spherical harmonic basis set to approximate the actual head shape [103]. The head is assumed to be a homogeneous single volume with an isotropic conductivity profile. The single shell model requires the specification of the number of harmonics to use: larger numbers of harmonics provide an improved representation of the head shape but come at an additional computational expense. Nolte et al. recommend 20 harmonics as sufficient for most practical applications but critically note that more eccentric sources have a higher error and so require more harmonics to achieve a given accuracy.

2.5 Inverse solutions and source reconstruction

Given an estimate of the lead-fields, we now turn our attention to inferring the underlying neural activity. From this point onwards we will assume that the error in the estimates of the MEG lead-fields are negligibly small (regardless of the estimation technique used).

Knowing the exact lead-fields is insufficient to solve the MEG inverse problem. We must make some additional assumptions to make the inverse problem tractable. There is a huge range of source reconstruction approaches.

2.5.1 Dipole fitting

The simplest constraint is to assume that the brain activity is generated by a small number of current dipoles. Under this assumption, one can then estimate the location,

orientation and time course of these dipoles that minimises the model error. However, dipole fitting approaches are limited to a few active sources and so are ill-suited to whole-brain connectivity analysis.

2.5.2 Distributed source approaches

Distributed source approaches attempt to simultaneously model the activity at every location of a grid spanning the whole brain. Given a user-specified grid, they impose some prior constraint for the solution of current dipoles at each vertex of the grid that makes the inverse problem tractable. One limitation of distributed approaches is their sensitivity to the grid coverage and density. If the grid excludes regions or has an insufficient density of vertices, then the resulting estimate of neural activity will be inaccurate.

Minimum Norm Estimates (MNE) are a popular class of distributed source reconstruction techniques [67]. Given a distributed grid of dipoles, usually spanning the cortex, MNE infers the optimum solution by L2-norm optimisation [70]. As we stated previously, the basic inverse problem, $\mathbf{X} = \mathbf{H^{-1}Y}$, is ill-posed because the number of neuronal sources making up the underlying brain activity, $\mathbf{X}$, greatly exceeds the number of sensors making up the MEG data $\mathbf{Y}$, such that the lead-field matrix, $\mathbf{H}$, is singular. MNE imposes an L2-norm regularisation which makes $\mathbf{H}$ invertible. This corresponds to including a minimum energy constraint for the estimated source solution. One consequence of the MNE formulation is that distributed source estimates tend to be preferentially smooth and superficial, which may not reflect the true underlying neuronal activity.

MNE does not have any interference rejection properties: artefacts in the data will be projected into the source-space solution. MNE is robust to correlated sources, unlike beamforming which is discussed later. MNE performs best when reconstructing spatially distributed activity but is poorer at reconstructing localised or deep brain activity [70]. In this study, we do not use MNE as a source reconstruction technique but note that it is used widely by other groups [43, 44, 90]. Preliminary research comparing minimum norm and beamforming suggests that they perform similarly in

MEG FC analyses [106].

Multiple sparse priors (MSP) is another class of distributed source reconstruction techniques [57]. MSP is a Bayesian framework for automatically and efficiently selecting a set of priors to impose on the source reconstruction that best maximise the model evidence. These priors can be sparse or distributed. Consequently, MSP has the advantage over MNE of being able to produce both sparse and distributed solutions depending on the what type of solution the MEG data best supports.

2.5.3 Beamforming

Beamforming is a source-reconstruction approach that does not directly attempt to solve the inverse problem but instead uses the forward model to design a spatial filter that estimates the activity at a location of interest.

For simplicity, we will assume that the orientation of each dipole is known and this information is incorporated into the lead-fields (i.e. $\mathbf{H}_q$ is now a N by 1 vector). However, the concepts described here are equally applicable to the full vector description of the neuronal dipoles. In section 2.5.3.4, we present a selection of techniques for estimating the dipole orientation.

In beamforming, we attempt to find an estimate of the neural activity, $\mathbf{x}_{q'}$, at location q', as a weighted sum of the data $\mathbf{Y}$.[2]

$$\hat{\mathbf{x}}_{q'} = \mathbf{W}_{q'}\mathbf{Y} \tag{2.3}$$

Ideally, our filter would perfectly preserve the activity at the location of interest q' but would completely reject any signals originating from other locations. This is not possible unless the number of active sources is very small. Instead, we design a spatial

[2]Notation: in the following section and subsequently through this thesis, the following notation will be adopted:

- The true underlying value of a variable has no accent: e.g. $\mathbf{x}_q$.
- The raw beamformer estimate of a variable is denoted with a circumflex: e.g. $\hat{\mathbf{x}}_q$.
- The *weights-normalised* beamformer estimate of a variable is denoted with a tilde: e.g. $\tilde{\mathbf{x}}_q$.

filter that minimises the total power passed whilst perfectly preserving the activity at the location of interest q'. Mathematically we can formulate this constraint as:

$$\min_{\mathbf{W}_{q'}}(trace(\mathbf{W}_{q'}\mathbf{C}\mathbf{W}_{q'}^{T})) \tag{2.4}$$

subject to

$$\mathbf{W}_{q'}\mathbf{H}_{q'} = 1 \tag{2.5}$$

which can be solved (using Lagrange multipliers) to give the following formulation for the beamformer weights:

$$\mathbf{W}_{q'} = [\mathbf{H}_{q'}^{T}\mathbf{C}^{-1}\mathbf{H}_{q'}]^{-1}\mathbf{H}_{q'}^{T}\mathbf{C}^{-1} \tag{2.6}$$

where $\mathbf{C}^{-1}$ is the inverse of the data covariance matrix $\mathbf{C} = \frac{1}{N_{samples}}\mathbf{Y}\mathbf{Y}^{T}$. In some cases, it is necessary to regularise the data covariance matrix, usually by boosting the diagonal:

$$\mathbf{C}_{regularised} = \mathbf{C} + \mu \times \lambda_{minimum} \times \mathbf{I} \tag{2.7}$$

where μ is a user-specified arbitrary scaling factor and $\lambda_{minimum}$ is the smallest eigenvalue of the data covariance and is an estimate of the variance of the noise [25]. Regularisation typically sacrifices spatial resolution but increases the temporal signal-to-noise ratio. Techniques exist to infer the optimum regularisation from the data, but have an associated computational cost [140].

2.5.3.1 Correlated sources

The beamformer has a major limitation that is a consequence of its formulation. In the presence of two (or more) highly correlated sources, the beamformer will fail to estimate the variance of the sources, regardless of their spatial separation. Instead,

it will blur the sources into a single mis-localised estimate. This is because the beamformer cancels out the shared contribution of the correlated sources when it estimates the set of weights that best minimise the total power projected at each voxel. Van Veen et al. demonstrated this point clearly using simulations of highly correlated sources [134]. Furthermore, Brookes et al. simulated this using a range of correlations, demonstrating that the effect becomes significant when the phase lag between two sources drops below 60 degrees [27].

Fortunately, empirical evidence suggests that stationary correlations (of sufficient strength to degrade beamformer performance) are not observed in resting-state MEG data. Multiple functional connectivity studies have successfully used the beamformer to detect plausible FC, suggesting that the non-stationarity and variability in the MEG signal (over several minutes) protects against correlated source cancellation [25,30,88]. Furthermore, recent work, which directly compares resting-state networks estimated using beamformers and minimum-norm solutions, finds minimal differences when utilising either method [106].

2.5.3.2 Advantages of beamforming

Beamforming has three major advantages as a source reconstruction technique. Firstly, from the definition of the beamformer weights (equation 2.6) at a given location, we can see that the weights are a function of the data covariance matrix and the lead-fields at that location. This means that the estimated solution at any location does not depend on the number of dipoles estimated, the coverage of the dipole grid nor the density of the grid used. This makes the beamformer extremely flexible for multi-scale analyses (from single voxel source reconstruction through to whole-brain estimates).

Secondly, the spatial specificity of the beamformer enables a degree of interference suppression. This is well demonstrated by Brookes et al. who show that the beamformer suppresses the widespread ECG artefact at the source-level [28]. The beamformer will intrinsically suppress any non-neuronal artefacts that do not originate from the location of interest.

Thirdly, the beamformer has reduced signal leakage between regions compared

with MNE techniques [118]. Signal leakage refers to the spurious correlations introduced between source-reconstructed estimates of two voxels. We discuss this in more detail in section 2.6.

2.5.3.3 Beamformer *weights normalisation*

In this section, we contrast the beamformer estimate of the variance of a dipole with the true underlying dipole variance. In doing so, we show that there is a spatially varying contribution to the beamformer variance which is a product of the projected sensor noise. We introduce *weights normalisation* as a correction that makes the noise projection uniform across the brain.

We can evaluate the following expression for the variance, $\hat{\sigma}^2_{q'}$, of our beamformer estimate of the neural activity as:

$$\hat{\sigma}^2_{q'} = \frac{1}{N_{samples}} \hat{\mathbf{x}}_{q'} \hat{\mathbf{x}}^T_{q'} \tag{2.8}$$

As we stated previously, we estimate the neural activity at q' as a weighted sum of the sensor data (described by equation 2.3), which in turn is a function of the true neural activity and the lead-fields (described by equation 2.2). Substituting equations 2.2 and 2.3 into equation 2.8, we get:

$$\hat{\sigma}^2_{q'} = \frac{1}{N_{samples}} \mathbf{W}_{q'} (\mathbf{H}_{q'} \mathbf{x}_{q'} + \sum_{q \neq q'} \mathbf{H}_q \mathbf{x}_q + \boldsymbol{\varepsilon}) (\mathbf{H}_{q'} \mathbf{x}_{q'} + \sum_{q \neq q'} \mathbf{H}_q \mathbf{x}_q + \boldsymbol{\varepsilon})^T \mathbf{W}^T_{q'} \tag{2.9}$$

If we substitute the unity pass band beamformer constraint ($\mathbf{W}_{q'} \mathbf{H}_{q'} = 1$) and make the following assumptions:

1. All the dipoles are uncorrelated with each other. This is a requirement for beamforming. In the case of long periods of resting-state data this is a reasonable assumption.

2. All the dipoles are uncorrelated with the sensor-level noise, $\boldsymbol{\varepsilon}$.

we can then express expected value of the variance of the beamformer estimate as:

$$\begin{aligned}\hat{\sigma}^2_{q'} &= \frac{1}{N_{samples}}(\mathbf{x}_{q'}\mathbf{x}_{q'}^T + \mathbf{W}_{q'}(\sum_{q\neq q'}\mathbf{H}_q\mathbf{x}_q\mathbf{x}_q^T\mathbf{H}_q^T)\mathbf{W}_{q'}^T + \mathbf{W}_{q'}\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}^T\mathbf{W}_{q'}^T) \quad (2.10)\\ &= \underbrace{{\sigma_{q'}}^2}_{\text{true variance}} + \underbrace{\mathbf{W}_{q'}(\sum_{q\neq q'}\mathbf{H}_q{\sigma_q}^2\mathbf{H}_q^T)\mathbf{W}_{q'}^T}_{\text{signal leakage}} + \underbrace{{\sigma_{\boldsymbol{\varepsilon}}}^2\mathbf{W}_{q'}\mathbf{W}_{q'}^T}_{\text{noise projection}}\end{aligned}$$

where we expressed the sensor-level noise covariance matrix as the noise variance ${\sigma_{\boldsymbol{\varepsilon}}}^2$ multiplied by an N by N identity matrix (where N is the number of sensors).

The sensor noise term in equation 2.10 has a spatially-varying bias ($\mathbf{W}_{q'}\mathbf{W}_{q'}^T$) which leads to an overestimation of the variance in the centre of the brain. This depth bias can be seen in the *raw* beamformer and envelope variance maps shown in Figure 2.4, taken from [65]. Usually, we divide our estimate, $\hat{\mathbf{x}}_{q'}$, by $\sqrt{\mathbf{W}_{q'}\mathbf{W}_{q'}^T}$. This down-weights deeper voxels and gives an unbiased estimate of the noise variance across all voxels. Figure 2.4 also shows the *weights-normalised* (or *scaled*) beamformer and envelope variances. Now the peak variance originates from the cortex. Without *weights normalisation*, we observe an increase in the variance of the estimated neural activity even if the true neural activity has homogeneous variance [134]. While *weights normalisation* corrects for this, it also creates a biased estimate of the true variance. This means that the *weights-normalised* beamformer estimate of a dipole's variance, $\tilde{\sigma}^2_{q'}$, is no longer equal to the true variance, $\sigma^2_{q'}$, plus some error contributions due to signal leakage and sensor noise. Instead, the beamformer estimate is equal to the true variance scaled by the *weights normalisation*, which we show mathematically below:

$$\begin{aligned}\tilde{\sigma}^2_{q'} &= \frac{\hat{\sigma}^2_{q'}}{\mathbf{W}_{q'}\mathbf{W}_{q'}^T} \quad (2.11)\\ &= \frac{{\sigma_{q'}}^2}{\mathbf{W}_{q'}\mathbf{W}_{q'}^T} + \frac{\mathbf{W}_{q'}(\sum_{q\neq q'}\mathbf{H}_q\sigma_q^2\mathbf{H}_q^T)\mathbf{W}_{q'}^T}{\mathbf{W}_{q'}\mathbf{W}_{q'}^T} + {\sigma_{\boldsymbol{\varepsilon}}}^2 \quad (2.12)\end{aligned}$$

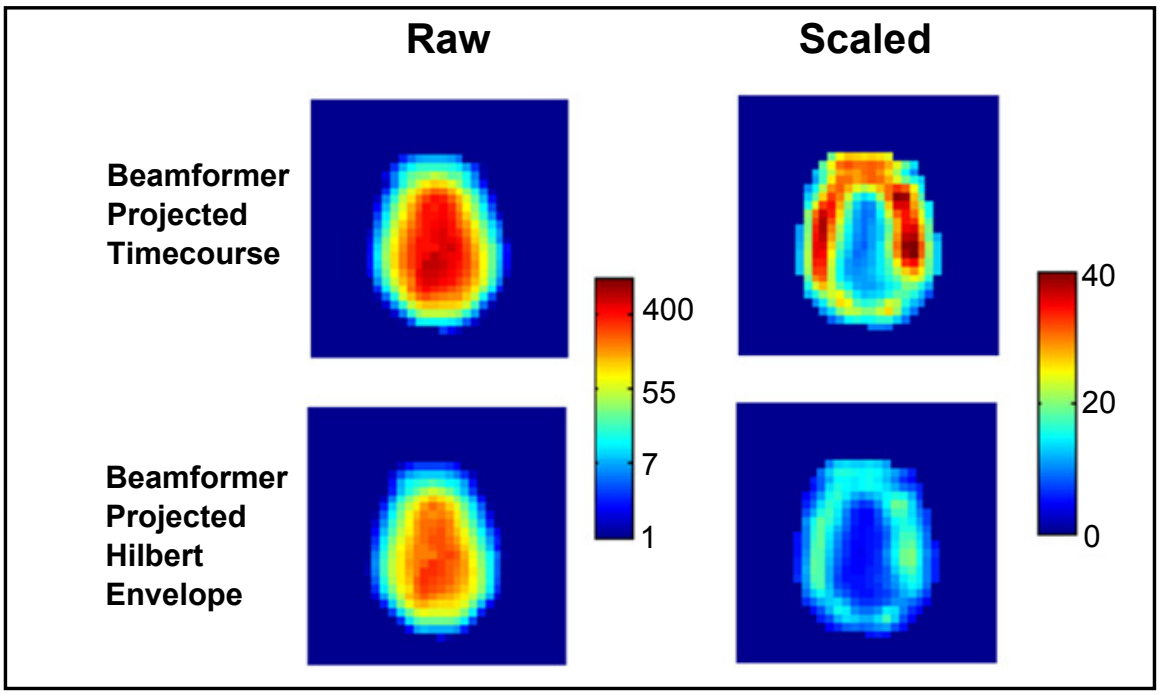


Figure 2.4: *An example map (axial slice) of the beamformer estimates of variance before (raw) and after (scaled) weights normalisation. The raw variance estimated greatly increases with depth. After weights normalisation, the peak variance now originates from the cortex. This effect applies both to the oscillatory time course and its amplitude envelope. Taken from [65, Fig. 2A].*

In chapter 5, we show that this bias in the estimate of the variance causes major problems when performing statistics over multiple beamformer sessions.

2.5.3.4 Estimating the dipole orientation

At the beginning of section 2.5.3, we assumed the dipole orientation at every vertex was known. In this section, we will elaborate on several techniques to estimate the orientation. Using our forward model, we can estimate three lead-field vectors, $\mathbf{H}_{q,i}$, $\mathbf{H}_{q,j}$ and $\mathbf{H}_{q,k}$ which equal the activity observed at the sensor array due to the i, j and k contributions of a unit current dipole at location q. We can combine these three vectors into a single lead-field matrix $\mathbf{H}_{q,i,j,k} = [\mathbf{H}_{q,i}, \mathbf{H}_{q,j}, \mathbf{H}_{q,k}]$.

Let us assume that the orientation of the current dipole at q is described by the 3 by 1 unit vector $\boldsymbol{\theta}_q$. We can project our full lead-field estimate, $\mathbf{H}_{q,i,j,k}$ onto $\boldsymbol{\theta}_q$ to estimate the rank-reduced lead-field $\mathbf{H}_q$, which corresponds to the activity observed at the sensors due to a unit dipole of orientation $\boldsymbol{\theta}_q$:

$$\mathbf{H}_q = \mathbf{H}_{q,i,j,k}\boldsymbol{\theta}_q \tag{2.13}$$

Often, the dipole orientation, $\boldsymbol{\theta}_q$, is constrained to the plane tangential to the head surface. This constraint reflects the lack of contribution of the radial component of a current dipole in a spherical conductor to the externally measured magnetic field [117].

Constraining the orientation to the tangential plane creates a one-dimensional problem where the unknown orientation now lies between 0 and 2π. We can further constrain the search space to between 0 and π, noting that orientations in the range π to 2π will give identical, except sign-flipped, time courses. This is known as the problem of *dipole sign ambiguity* and is a consequence of two dipoles of opposing orientations being equally good estimates of the underlying neural activity.

We will now present three techniques for estimating the full orientation of each source-reconstructed dipole.

Anatomically constrained dipole orientations: In section 2.1.3, we outlined the hypothesis that the MEG signal originates from the synchronous activity of thousands of neurons which are aligned perpendicularly to the cortical sheet. Based on this reasoning, one approach to estimating the source orientations is to constrain them to be perpendicular to the subject's cortical surface (which can be extracted from their structural MRI) [7, 104]. This technique requires the source reconstruction to be confined to the cortical surface and hence overlooks potential contributions from deep brain sources. Furthermore, it requires a structural MRI for each participant and an accurate cortical surface extraction. This may not be available (for example with participants with non-MR compatible medical implants).

Maximising projected power: The second technique for estimating the orientation is to use the orientation that maximises the projected power of the dipole to the sensor array. This can be estimated as the first principle component of the vector beamformer solution [119]. We utilise this technique as it provides a balance between adaptability (i.e. it can be applied to any voxel and it does not require a structural

MRI) and computational simplicity and speed.

Maximising signal-to-noise ratio - Synthetic Aperture Magnetometry: Synthetic Aperture Magnetometry (SAM) beamformers estimate the optimum orientation of a dipole at the q^{th} location by performing a search for the orientation that maximises the signal-to-noise ratio [115, 136]. The signal-to-noise ratio is measured using the pseudo-Z-statistic, $\mathcal{Z}_q$, which is defined as:

$$\mathcal{Z}_q = \frac{\hat{\sigma}_q^2}{{\sigma_\varepsilon}^2} \tag{2.14}$$

SAM estimates $\mathcal{Z}_q$ for tangentially constrained dipole orientations between 0 and π and uses the orientation that gives the maximum pseudo-Z-statistic as the optimal estimate.

2.6 Artificial functional connectivity in MEG

In MEG (and EEG), FC measurements are plagued by artificial contributions to connectivity metrics [118, 138]. Consider a case where we have a grid of mutually uncorrelated dipoles. Each MEG sensor's recording will be a linear mixture of these dipole time courses. As such all the sensors will have some degree of correlation with each other - we term this contribution *field spread* [138]. Measuring functional connectivity between sensors is hugely confounded by the field spread artefact.

Projecting the data back into source-space has been shown to reduce the artificial FC observed [118]. However, it does not eliminate it. As stated in section 2.4, source-reconstructed dipoles are weighted sums of the sensor data (equation 2.3) Therefore, two reconstructed dipoles will have correlated activity if their weights are correlated [25, 65]. If we combine equations 2.2 and 2.3, we can mathematically show that our estimate for one voxel contains activity *leaked* from all others:

$$\hat{\mathbf{x}}_{q'} = \mathbf{W}_{q'}(\mathbf{H}_{q'}\mathbf{x}_{q'} + \underbrace{\sum_{q \neq q'} \mathbf{H}_q \mathbf{x}_q}_{\text{signal leakage}} + \boldsymbol{\varepsilon}) \tag{2.15}$$

This source-space contribution to FC is referred to as *signal leakage* or *cross-talk*. It is possible to eliminate signal leakage in certain cases (such as using multi-dipole beamformers to estimate pairwise FC [98]). Generally, accounting for signal leakage contributions to FC is a significant challenge in whole-brain FC analysis. We discuss practical solutions to account for signal leakage in section 4.4.1.

2.7 MEG systems

In this thesis, data acquired on two different MEG systems are used. Although, theoretically the scanners measure similar signals and are designed using similar assumptions, technical differences in the design and initial analysis of these systems require subtle variations in the subsequent analysis of their outputs. Below is a brief introduction to the two systems utilised in this thesis.

2.7.1 *Elekta Neuromag*

The *Elekta Neuromag* has an array of 102 magnetometers and 204 planar gradiometers (102 orthogonal pairs). By overlapping each magnetometer and pair of gradiometers, this array gives three complementary measurements at 102 locations. This array can detect fields as low as 5fT [49]. However, correctly combining the two sensor types is not trivial and is discussed in section 3.3.1.

The *Elekta Neuromag* has an interference rejection algorithm that employs signal space separation (SSS) [127]. Signal space separation projects the MEG data into a basis set of spherical harmonics such that sources from within the sensor helmet can be separated from any sources of interference that do not originate from within the helmet. This is implemented using *MaxFilter*^{TM}, a software package provided with the *Elekta Neuromag* system [50]. *MaxFilter*^{TM} uses an implementation of SSS to attenuate any sources outside the helmet such as electromagnetic cardiac interference. However *MaxFilter*^{TM} has two technical caveats: firstly, it greatly reduces the dimensionality of the MEG data. Secondly, *MaxFilter*^{TM} can propagate certain

artefacts from an individual channel into multiple channels. We discuss these issues and our solutions in chapter 3.

MaxFilter^{TM} can also be used to apply movement compensation. When the MEG data is projected into the spherical harmonic basis set, it can be re-projected into an arbitrary location. By continuously measuring the head position throughout the scan, *MaxFilter*^{TM} can re-project the data at each time point into a specific head position.

2.7.2 *275 channel CTF whole-head system*

The *275 channel CTF whole-head system* consists of an array of 275 axial gradiometers, which are at different locations (giving a greater density of sensors but with only one type of measurement being recorded at each location). In addition, within the magnetically shielded room there is an array of 25 reference sensors. This reference array is designed to measure the same far-field interference as the scanner helmet but is placed sufficiently far from the subject that it is insensitive to any neural activity. The recordings from the reference array are projected onto each of the 275 axial gradiometers and then removed. This is referred to as **3rd order gradiometer correction** and is very effective at rendering the main sensor array insensitive to far-field interference [136].

2.8 Summary

This chapter represents a brief and far from exhaustive introduction to MEG with an extended focus on source reconstruction using beamformers. We have laid the mathematical framework that will be used in subsequent chapters where we address practical challenges to MEG analysis and present a framework for performing functional connectivity analyses in resting-state and task-positive MEG.

Chapter 3

Preprocessing for MEG analyses

3.1 Chapter abstract

In this chapter, we outline the MEG preprocessing framework that enables us to perform the functional connectivity analyses used throughout this thesis. We adopt a MEG preprocessing pipeline with three major novel contributions. The first is a procedure for applying *MaxFilter*^{TM} to *Elekta Neuromag* data that is robust to corruption from specific channel artefacts. The second is an implementation of independent component analysis for de-noising multi-sensor-type MEG data in sensor-space. The third is a set of modifications to the LCMV beamformer that enables source reconstruction of multi-sensor-type **and** *MaxFiltered* MEG data. We have incorporated these insights into our standard preprocessing pipeline (summarised in Figure 3.7).

3.2 Preprocessing with *MaxFilter*^{TM}

Signal space separation can be used to suppress far-field sources of interference [127]. We use the commercial implementation of *MaxFilter*^{TM} with its standard settings as it has been optimised for use with the *Elekta Neuromag* system [50].[1]

However, we have identified a potential issue when applying *MaxFilter*^{TM} to *Elekta*

[1]The signal space separation algorithm has not been implemented for other scanners, such as the *CTF* 275-channel system. Instead, data acquired on the *CTF* scanner, also used in this thesis, has 3rd order gradiometer correction which is a combined hardware/software technique for suppressing far-field interference [136].

Neuromag data. Specific electronic artefacts[2] introduced in the MEG recording get propagated from one channel into every channel by the *MaxFilter*^{TM} algorithm. These artefacts are typically characterised by high amplitude, sharp discontinuities. We demonstrate this in Figure 3.1. The channel artefact has been recorded in a single magnetometer (left panel of Figure 3.1A). If this channel is included in the *MaxFilter*^{TM} analysis, then the artefact is propagated into every channel (Figure 3.1B). However, if the channel is excluded from the *MaxFilter*^{TM} then the resulting data are not corrupted (Figure 3.1C). One option is to forgo using *MaxFilter*^{TM} completely. Although this protects against the propagation of scanner electronics artefacts, it has two shortcomings: firstly, *MaxFilter*^{TM} can improve the quality of the data. We demonstrate this by identifying two noisy gradiometers in Figure 3.1A which are cleaned up with *MaxFilter*^{TM}. Secondly, *MaxFilter*^{TM} enables the use of continuous motion correction which may improve the sensitivity of some analyses (particularly if the participants move a lot, such as when scanning children).

To account for channel artefacts we implement a two stage procedure. In the first stage, the data are visually inspected and any channels containing high amplitude, sharp discontinuities are flagged. In the second stage, *MaxFilter*^{TM} is applied but ignores any flagged channels. The data are visually inspected to confirm that no artefacts have been propagated across channels. We have found this to be a simple but robust approach.

3.3 De-noising with ICA

Independent component analysis (ICA) is used routinely to separate physiological artefacts from MEG and EEG data [91]. It can be used to remove artefacts due to cardiac activity, eye-blinks and saccades, muscle activity and mains interference.

[2]These artefacts are a consequence of the *Elekta Neuromag* scanner's automatic tuning, which tries to minimise the noise levels across the sensors. We have found that the tuning occasionally selects an unstable minimum-noise operating point where the relationship between the flux passing through each **SQUID** and the voltage measured across that **SQUID** can transiently move from an approximately linear regime into a highly non-linear regime. These transient regime shifts result in a sharp spike in the sensor time course, followed by a decay of several seconds back to the normal operating regime.

Numerous ICA de-noising tools are available in open source software packages (e.g. **FieldTrip** (*Donders Institute for Brain, Cognition and Behaviour, Radboud University*)). However, these tools can only de-noise single-sensor-type data. However, as we shall see, particular care is needed when applying ICA de-noising to multi-sensor-type data, such as that acquired using an *Elekta Neuromag* scanner (which has both magnetometer and planar gradiometer sensors). Here, we develop an ICA de-noising framework specifically designed to de-noise *Elekta Neuromag* multi-sensor-type data. Our ICA de-noising framework has 4 stages:

1. Sensor normalisation
2. ICA decomposition
3. Identification of artefact components
4. Removal of artefact components

3.3.1 Sensor normalisation

Elekta Neuromag data has two sensor types: magnetometers and planar gradiometers. The two sensor types have different units and different noise levels. Many ICA implementations start with a dimensionality reduction, typically via a principle component analysis (PCA). The PCA removes the smaller principle components and the ICA is then performed on the reduced dimensionality subspace. In single-sensor-type analyses, this is straightforward, based on the idea that the artefacts we seek often have greater variances than the signals due to neural activity (refer to Figure 2.3), and so low dimensional PCAs focus on these components.

However, magnetometers have a much smaller signal and noise variance compared with planar gradiometers (Figure 3.2A). The spectrum of eigenvalues has two distinct regimes corresponding to the planar gradiometers and magnetometers. The corresponding eigenvectors show that most of the magnetometer information is confined to the 102 smallest principle components. As such, dimensionality reductions to

dimensions less than 204 will remove nearly all the magnetometer information. Consequently, the ICA effectively operates only on the gradiometers and only de-noises them.

In principle, one could avoid the issue of sensor-normalisation by performing ICAs separately on the different sensor types. However, this will lead to each sensor type being de-noised differently which in turn will cause downstream issues in the source reconstruction.

Therefore, in order to perform ICA correctly on multi-sensor-type data, it is essential to normalise the different sensor types. One possible approach is to normalise each sensor type separately to have unit average variance [72]. Figure 3.2B shows the spectrum of eigenvalues and corresponding eigenvectors for variance-normalised magnetometers and gradiometers. It is clear that variance normalisation does not lead to both sensor types being equally well represented in the largest principle components. This is because the eigenvalues of the magnetometers span a much greater range (several orders of magnitude greater) than the eigenvalues of the gradiometers (this can be seen in the eigenspectrum in Figure 3.2A). Variance normalisation approximately matches the largest eigenvalues of the two sensor types, but the remaining eigenvalues of the gradiometers are still much larger than the equivalent eigenvalues of the magnetometers.

Our proposed solution is to "noise-normalise" each sensor type to have equal noise variances. Noise variances could be estimated from empty room MEG data but this requires an empty room sample to be acquired with each session of data [72]. Instead, we use the smallest eigenvalue of each class of sensor as an estimate of its noise variance. Figure 3.2C shows the same data after minimum eigenvalue normalisation. Now information from both sensor types is present in the largest principle components. Consequently, both sensor types will be de-noised equally. This is in agreement with Beckmann et al.'s (2004) implementation of ICA in the context of fMRI, where the data is noise-normalised across voxels [15].

When applying the minimum eigenvalue normalisation to *Elekta Neuromag* data that has had signal-space-separation applied to it, we have to modify this procedure to

account for the reduced rank. The true rank of the *MaxFiltered* data is approximately 64 but numerical rank estimators in Matlab (including the singular value decomposition used when estimating pseudo-inverses) often over predict this (see Figure 3.5A for a demonstration of this in the context of beamforming). As such we force the normalisation to use the 60th eigenvalue (the exact dimensionality cut-off is not precisely defined and so for stability we use the fifth smallest eigenvalue).

3.3.2 ICA decomposition

The sensor-space MEG data, $\mathbf{Y}$, is noise-normalised (using the minimum eigenvalue normalisation method described in the previous section) to give $\bar{\mathbf{Y}}$ which is in turn decomposed into a set number of I temporally independent components, $\mathbf{G}$ (an I components by T samples matrix), and associated sensor topographies, $\mathbf{F}$ (an N sensors by I components matrix), using **FastICA** [77]:

$$\bar{\mathbf{Y}} = \mathbf{FG} \tag{3.1}$$

The number of components is limited by the rank of the data but otherwise needs to be specified by the user. For non-*MaxFiltered Elekta Neuromag* data and *CTF* data, we estimate 150 components (we use a model order of 150 based on the preprocessing pipeline described by Mantini et al. [90]). In the case of *MaxFiltered* data, the rank is 64 and therefore we decompose the data into 64 components.

3.3.3 Identification of artefact components

For certain data sets, ECG, EOG and optical eye-tracker data were acquired with the MEG. For these data sets, we estimated the Pearson correlation coefficient between each component time course and the raw time course of each external signal. We classified any component with a correlation coefficient greater than 0.3 as an artefact. These data sets validated our procedure for classifying artefacts in data sets with no ECG, EOG or eye-tracker available.

We classified artefact components from such data sets via the following method.

Eye-blink, cardiac and mains interference components were manually identified by the combined inspection of the spatial topography, time course, kurtosis of the time course and frequency spectrum for all components. Eye-blink artefacts typically exhibit high kurtosis (>20), a repeated blink structure in the time course and very structured spatial topographies [53, 91]. Cardiac component time courses strongly resemble the typical ECG signals, as well as having high kurtosis (>20). Mains interference has extremely low kurtosis (typically <-1) and a frequency spectrum dominated by 50Hz line noise. Figure 3.3 shows the typical topographies, time courses and correlation values used for a single subject being de-noised with ICA.

3.3.4 Removal of artefact components

Having identified any artefact components, there are two options for estimating the de-noised MEG data. The first is to estimate it as the product of the non-artefact sensor topographies and independent time courses (i.e. to reverse the ICA decomposition excluding the artefact components).

$$\mathbf{Y}_{de-noised} = \mathbf{F}_{non-artefact}\mathbf{G}_{non-artefact} \tag{3.2}$$

The second approach is to estimate the de-noised sensor data as the subtraction of the artefact components from the raw data, shown below. This is equivalent to regressing out the artefact component time courses from the original data (where the regression parameters have been found as part of the ICA). As shown by Mantini et al. (2011), the subtraction is the preferred approach because we preserve any neural activity that might have been excluded from the ICA at the dimensionality stage [90].

$$\mathbf{Y}_{de-noised} = \mathbf{Y} - \mathbf{F}_{artefact}\mathbf{G}_{artefact} \tag{3.3}$$

We converted this subtraction into a multiplication using a residual forming matrix:

$$\mathbf{Y}_{de-noised} = (\mathbf{I} - \mathbf{F}_{artefact}\mathbf{G}_{artefact}\mathbf{Y}^{+})\mathbf{Y} \tag{3.4}$$

where $\mathbf{Y}^{+}$ is the pseudo-inverse of the sensor data $\mathbf{Y}$. This enabled us to simultaneously de-noise the data and correct the lead-fields via the montage function in **SPM8** (*FIL, UCL*).

In order to demonstrate the efficacy of our ICA de-noising pipeline, Figure 3.4 shows a single session of *Elekta Neuromag* data before and after ICA de-noising. Figure 3.4A shows 30s from the channel which has the strongest correlation with the eye-tracker blink time course. Note that several eye-blinks present in the raw data (blue) have been removed by the ICA. Furthermore, the channel's correlation with the eye-tracker blink time course has been reduced from 0.38 to -0.05. Figure 3.4B shows 10s of data from the channel with the strongest correlation with the ECG recording. The strong ECG artefact present in the raw data (blue) has been removed in the de-noised data (red). The overall channel correlation with the ECG has been reduced from 0.32 to -0.09. Finally, Figure 3.4C shows the reduction in 50Hz power in the mean of the magnetometer channels after ICA de-noising.

3.4 Removal of bad channels and bad epochs

In addition to *MaxFilter*^{TM} and ICA de-noising, we also perform a visual inspection of the sensor-space data to identify any remaining noisy channels and noisy epochs of data (these epochs can be of arbitrary length). Bad channels are excluded from the remaining analysis. Bad epochs are kept but excluded from specific stages (such as estimating correlation and covariance values) but are included for any operations which require the preservation of data continuity (such as temporal filtering or the Hilbert transform) [9].

3.5 Source reconstruction with a beamformer

In section 2.5.3 we introduced beamforming as a tool for estimating the source-level activity that drives MEG data [115, 134]. In this section, we introduce two novel

adaptations to generic beamformer implementations. The first adaptation is the manual specification of the rank of *MaxFiltered* data to ensure correct inversion of the data covariance matrix leading to robust beamforming of *MaxFiltered* data. The second is the application of minimum eigenvalue normalisation to allow beamforming of multi-sensor-type MEG data (such as *Elekta Neuromag* data).

3.5.1 Specifying the rank of *MaxFiltered* data to ensure robust inversion of the data covariance matrix

One consequence of applying $MaxFilter^{TM}$ is that the algorithm reduces the dimensionality of the data from 306 to 64. This reduction in rank reflects the removal of far-field contributions to the data. As such, the true dimension of the data covariance matrix is 64, as shown by Figure 3.5A which shows the eigenspectrum for a single subject's MaxFiltered data. As shown in equation 2.6, the beamformer weights are a function of the inverse of the data covariance matrix.

We found that the Matlab implementation of the pseudo-inverse consistently over predicts the rank of the data covariance matrix and as a consequence produces a poor estimate of its inverse. Figure 3.5A shows that the automatically estimated rank of the data covariance matrix is 68 which is after the knee of the eigenspectrum (see the red cross-hair). We demonstrated the consequence of this rank overestimation by performing a single session functional connectivity analysis. We beamformed a single session of *MaxFiltered* resting-state data, band-pass filtered into the β-band (13-30Hz) using this erroneous inverse covariance matrix. At every voxel of a 6mm grid, we estimated the correlation of the low-pass filtered ($<$0.5Hz), orthogonalised envelopes with a seed in the left motor cortex (MNI coordinates $[42, -24, 60]$mm). This functional connectivity analysis is described fully in section 4.4.2. Figure 3.5B shows the resulting correlation map: the strongest FC with the left motor cortex was found in the left frontal lobe rather than in any of the contralateral motor areas (which should show the strongest FC) [25, 74].

To fix this issue, we forced the pseudo-inverse to use a rank estimate that is smaller

than 64. Specifically, we forced the rank estimate to be 60 which is before the knee in the eigenspectrum (see the black cross-hair in Figure 3.5A). Figure 3.5C shows the equivalent seed-based correlation analysis: we found the strongest correlations with the left motor cortex to be in the medial and right motor areas, exactly where we would predict them to be.

When manually forcing the rank, it is also necessary to account for any reduced dimensionality introduced by the ICA de-noising. For all of our analyses on *MaxFiltered* data, we forced the rank of the inverse covariance matrix to be 60 to account for the combined reduction in rank caused by *MaxFilter*^{TM} and ICA de-noising.

3.5.2 Minimum eigenvalue value normalisation of multi-sensor-type MEG data

In section 3.3.1, we developed minimum eigenvalue normalisation in the context of performing ICA on sensor-space MEG data to remove physiological artefacts. However, it is also necessary to perform similar normalisation as a preprocessing step when beamforming over multiple sensor types of *MaxFiltered* data for the following reason. As we stated in the previous section, when beamforming *MaxFiltered* data, we must force the rank of the data to be 64 or less prior to estimating the inverse covariance matrix. This is equivalent to performing a PCA dimensionality reduction on the data covariance matrix. However, performing a dimensionality reduction (down to less than 64) on *MaxFiltered* data without any normalisation between sensor types will lead to the contribution from the magnetometers being excluded (in exactly the same way as we showed for sensor-space ICA de-noising). This can be seen in Figure 3.6A, which shows the eigenvalues/eigenvectors of a single session of *MaxFiltered* resting-state data without any normalisation. The eigenvectors show that the 64 largest principle components contain minimal information from the magnetometers. As such, it is essential to normalise each sensor type prior to the PCA dimensionality reduction. Figure 3.6 shows the same data after minimum eigenvalue noise-normalisation. Now, the 64 largest principle components contain information

from both the magnetometers and planar gradiometers.

There is an additional benefit to applying noise-normalisation to multi-sensor-type MEG data. In equation 3.5, we show an expression for the beamformer-estimate of the variance of a dipole in terms of the true variance of that dipole, the signal leakage from other dipoles, and the projection of the sensor noise.

$$\hat{\sigma}^2_{q'} = \frac{1}{N_{samples}}(\mathbf{x}_{q'}\mathbf{x}^T_{q'} + \mathbf{W}_{q'}\sum_{q\neq q'}\mathbf{H}_q{\sigma_q}^2\mathbf{H}^T_q\mathbf{W}^T_{q'} + \mathbf{W}_{q'}\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}^T\mathbf{W}^T_{q'}) \quad (3.5)$$

$$\hat{\sigma}^2_{q'} = {\sigma_{q'}}^2 + \mathbf{W}_{q'}\sum_{q\neq q'}\mathbf{H}_q{\sigma_q}^2\mathbf{H}^T_q\mathbf{W}^T_{q'} + {\sigma_{\boldsymbol{\varepsilon}}}^2\mathbf{W}_{q'}\mathbf{W}^T_{q'} \quad (3.6)$$

In this formulation, we assumed the sensor-level noise covariance matrix $\boldsymbol{\Sigma} = \frac{1}{T}\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}^T$ could be expressed as ${\sigma_{\boldsymbol{\varepsilon}}}^2\mathbf{I}$, essentially assuming that the noise across all the sensors has uniform variance and is uncorrelated. This led to the definition of the *weights normalisation* $\sqrt{\mathbf{W}_{q'}\mathbf{W}^T_{q'}}$ in section 2.5.3.3. In single-sensor-type MEG data (such as that acquired on a *CTF* system), this is a reasonable assumption as all the sensors are identical. However, in multi-sensor-type data this is not the case. In *Elekta Neuromag* data, the magnetometers and planar gradiometers will have different noise variances. Minimum eigenvalue normalisation is a data driven method for matching the noise levels of the two sensor classes. Without matching the noise levels, our *weights normalisation* must account for the different noise levels of the two channels and therefore must include the full noise covariance matrix. The *weights normalisation* term becomes $\sqrt{\mathbf{W}_{q'}\boldsymbol{\Sigma}\mathbf{W}^T_{q'}}$. Estimating the noise covariance matrix, $\boldsymbol{\Sigma}$, is not trivial - the covariance of empty room data could be used as a surrogate but this reflects both sensor noise and environmental interference. Also, such a method requires the acquisition of empty room MEG data for each scan because the sensor noise is not guaranteed to be stable over long periods of time.

3.6 The preprocessing pipeline

Here, we summarise our preprocessing pipeline which takes raw MEG data, de-noises it and performs a beamformer source reconstruction. Figure 3.7 represents our current best-practice procedure for preprocessing MEG data prior to any functional connectivity analysis and is used throughout this thesis. Any deviations from this pipeline will be explicitly stated in the relevant methods section. Our pipeline has the following stages:

- Data acquisition (either on an *Elekta Neuromag* or *CTF 275-channel whole head system*).
- Far-field interference suppression: *Elekta Neuromag - MaxFilter^{TM}. CTF 275-channel whole-head system - 3rd order gradiometer correction.*
- ICA de-noising to remove eye-blink, cardiac and 50Hz mains interference.
- Bad channel/bad epoch removal.
- Band-pass filtering.
- Co-registration of head shape to subject's structural using **SPM8** (*FIL,UCL*).
- Forward model estimation either using overlapping local spheres [75] or single shell forward models [103] implemented in **FieldTrip** (*Donders Institute for Brain, Cognition and Behaviour, Radboud University*).
- Beamformer source reconstruction to a 6mm grid spanning the whole head - including sensor normalisation and rank specification.

3.7 Summary

This chapter discussed the preprocessing pipeline used throughout this thesis. We focused on several critical pitfalls that one encounters when analysing *Elekta Neuromag* data. These challenges are consequences of the *MaxFilter^{TM}* algorithm and/or the

existence of multiple types of sensor. *MaxFilter*^{TM} has two critical pitfalls: the first is that it can propagate certain artefacts from one channel into many. In the best case, this leads to large numbers of channels or trials being discarded. In the worst case, the propagated artefact is overlooked and can lead to spurious results such as high functional connectivity. It is worth noting that the original manifestations of this artefact that led to its discovery were source-space independent components with whole-brain spatial maps.

Failing to account for the reduced rank of *MaxFiltered* data can be much more apparent as the beamformer gives very poor results, assuming that the analysis being performed has some ground truth (such as the known β-band FC between motor regions). Such test cases are essential for validating any analysis pipeline.

In this chapter, we also addressed the problem of combining multiple sensor types when performing ICA de-noising and beamforming. We focused exclusively on the fusion of magnetometers and planar gradiometers. We have found minimum eigenvalue normalisation to be a principled technique which gives empirically better results than any other technique. A future extension of minimum eigenvalue normalisation would be in the context of source-reconstructing concurrent EEG/MEG data. As EEG and MEG offer complementary insights and have different spatial sensitivity profiles (EEG is preferentially sensitive to sources in the gyri and sulci, whereas MEG is preferentially sensitive to sources between the gyri and sulci) a combined source reconstruction has the potential to offer superior results. The challenge of combining EEG/MEG in a single source-reconstruction solution is not trivial and is one of the hurdles preventing a wider adoption of the technique.[3]

The consequences of overlooking the issues outlined in this chapter can be quite insidious and only become apparent far downstream in any given analysis, whether it be a resting-state functional connectivity analysis or a task-based analysis.

[3]The complexity of EEG forward modelling and the additional time cost of setting up participants for combined EEG/MEG are also factors.

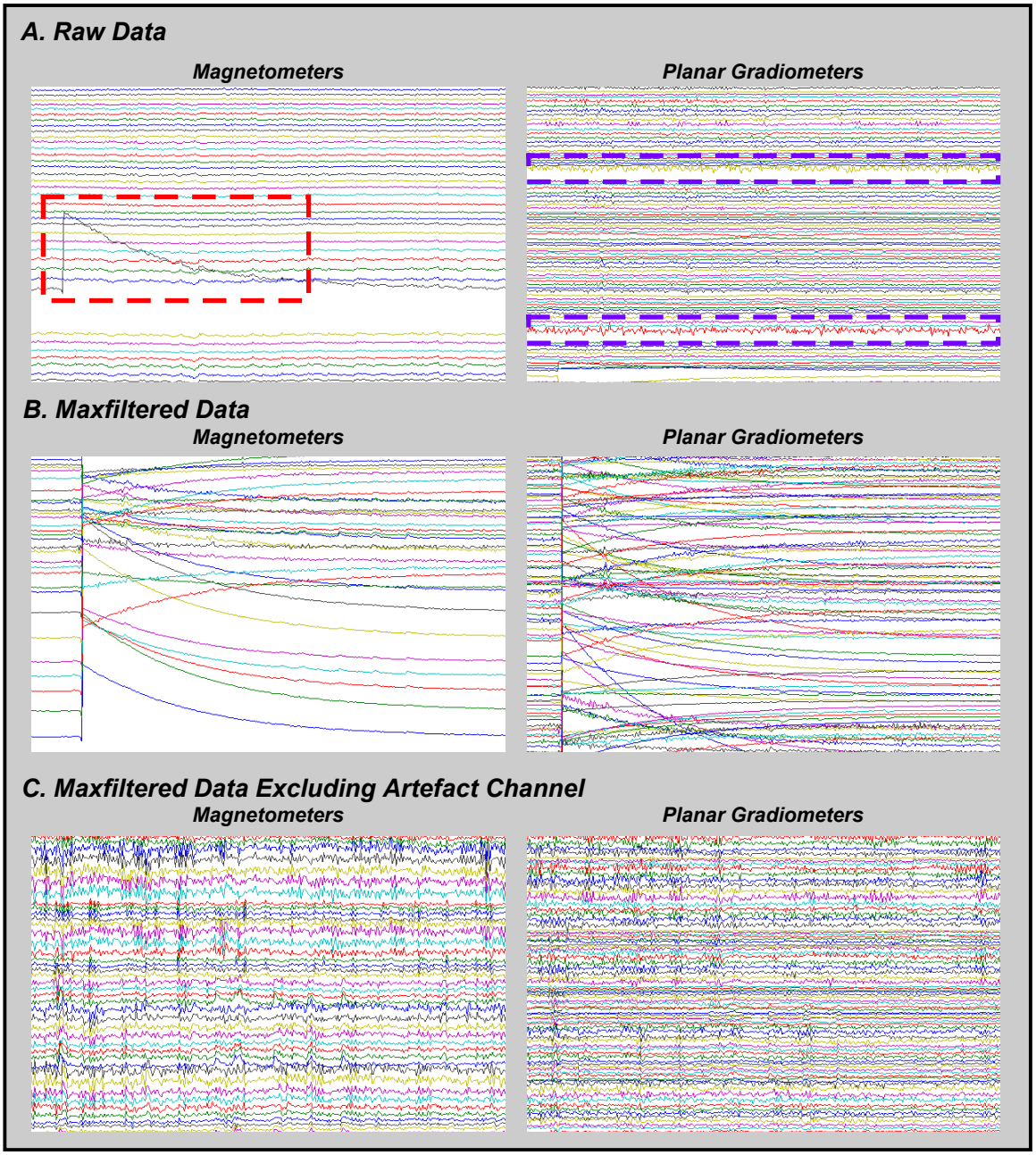


Figure 3.1: *Ten seconds of resting-state MEG data acquired on an Elekta Neuromag for a subset of channels. Magnetometers are shown on the left panels. Planar gradiometers are shown on the right. A. The raw data with no preprocessing applied. Note that a single magnetometer channel has a high amplitude channel artefact shown in the dashed red box. Two noisy planar gradiometers have also been identified and are highlighted with dashed blue boxes. B. The same data after applying MaxFilter^{TM}. The MaxFilter^{TM} algorithm has propagated the magnetometer channel artefact across all the channels. C. The same data after MaxFilter^{TM} has been applied, except that the corrupted magnetometer channel has been excluded. Note that the two noisy gradiometers have been successfully cleaned up by the MaxFilter^{TM} process.*

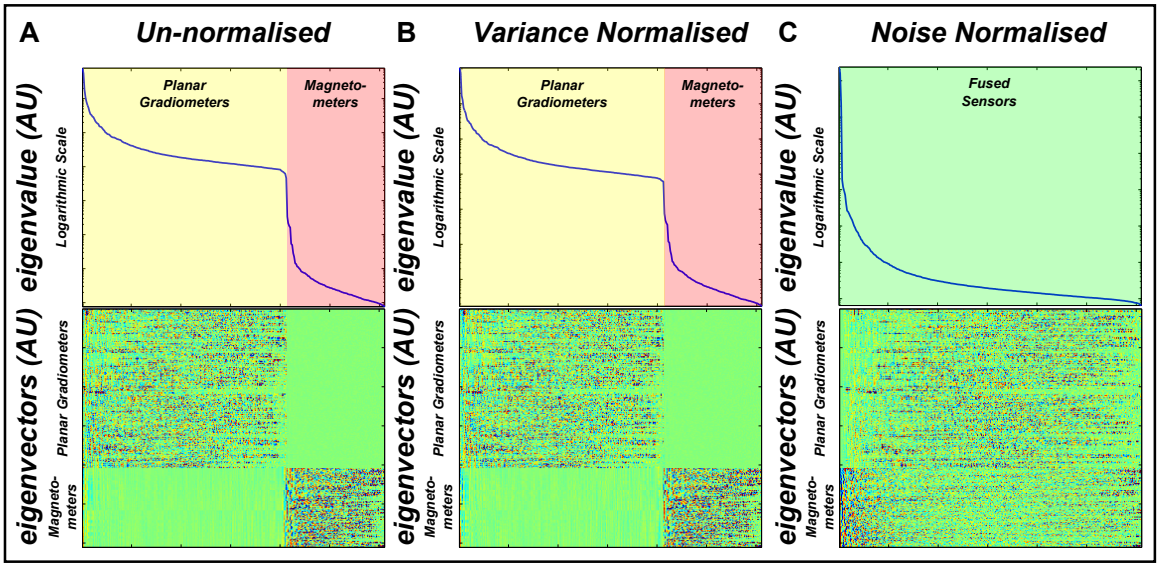


Figure 3.2: *Demonstration of minimum eigenvalue normalisation. A. The eigenspectrum and corresponding eigenvectors for a single session of Elekta Neuromag data. Each eigenvector is a column vector, where channels have been grouped by sensor type - 204 planar gradiometers followed by 102 magnetometers. Magnetometers have much smaller eigenvalues and, as such, performing a dimensionality reduction will exclude the signal from the magnetometers. B. The eigenspectrum and eigenvectors of the same session after each sensor type has been normalised by its average variance. Dimensionality reduction will still exclude most of the information contained in the magnetometers. C. The eigenspectrum and eigenvectors of the same session after each modality has been noise-normalised by its smallest eigenvalue. Minimum eigenvalue normalisation matches the noise levels of both sensor types and means that any arbitrary dimensionality reduction will include information from both magnetometers and planar gradiometers.*

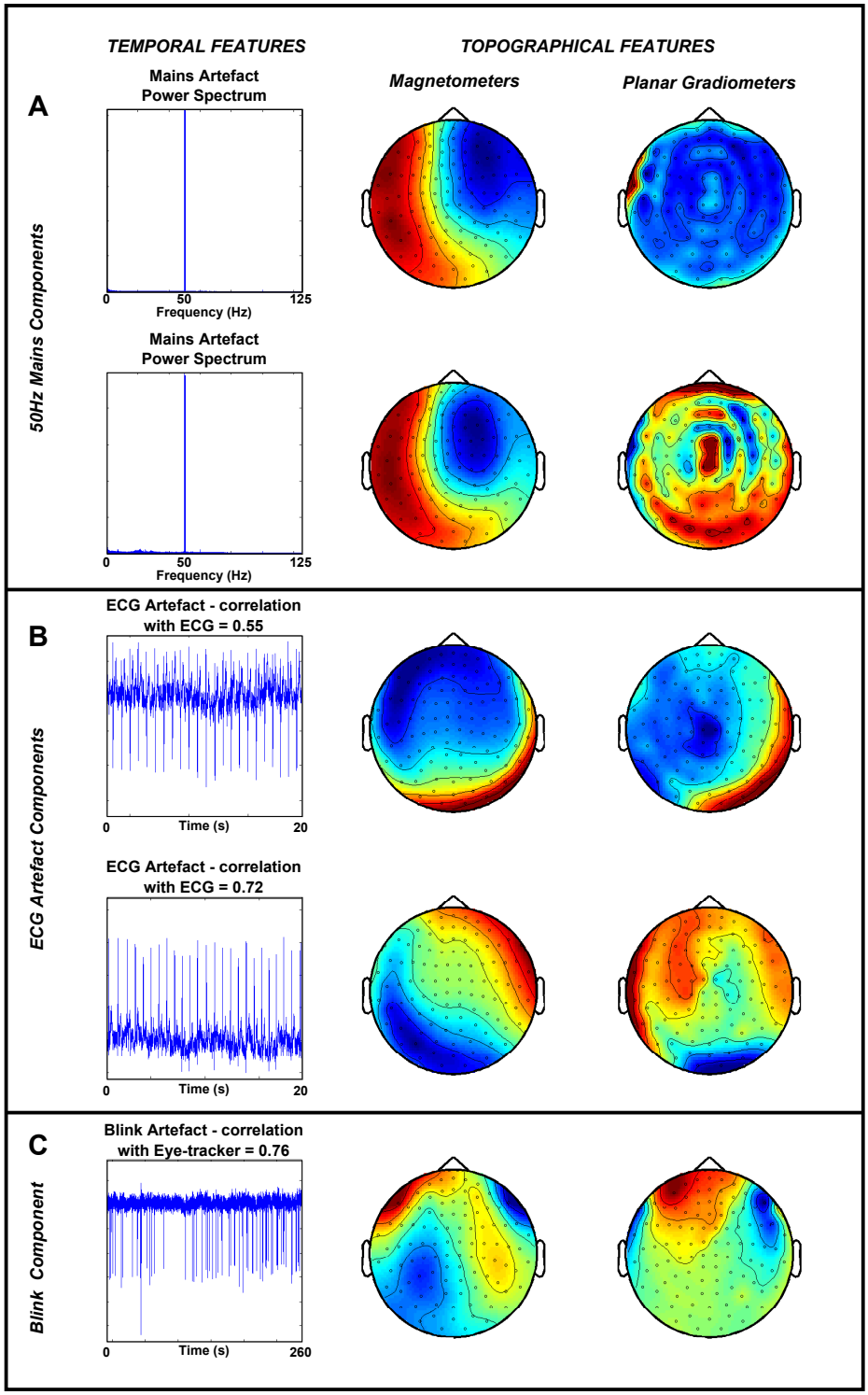


Figure 3.3: *An example of the temporal and topographical features used to manually classify artefact components in ICA de-noising (taken from a single session of Elekta Neuromag resting-state data). A. Two 50Hz mains components whose power spectra are dominated by 50Hz signals. B. Two cardiac components that have strong correlations with the ECG. The component time courses strongly resemble a typical ECG trace. C. A single eye-blink component that has strong correlations with the eye-tracker. The time course has regular blink events and the topographies of the magnetometers/gradiometers have symmetric maps with peak values over the frontal sensors, typical of an eye-blink.*

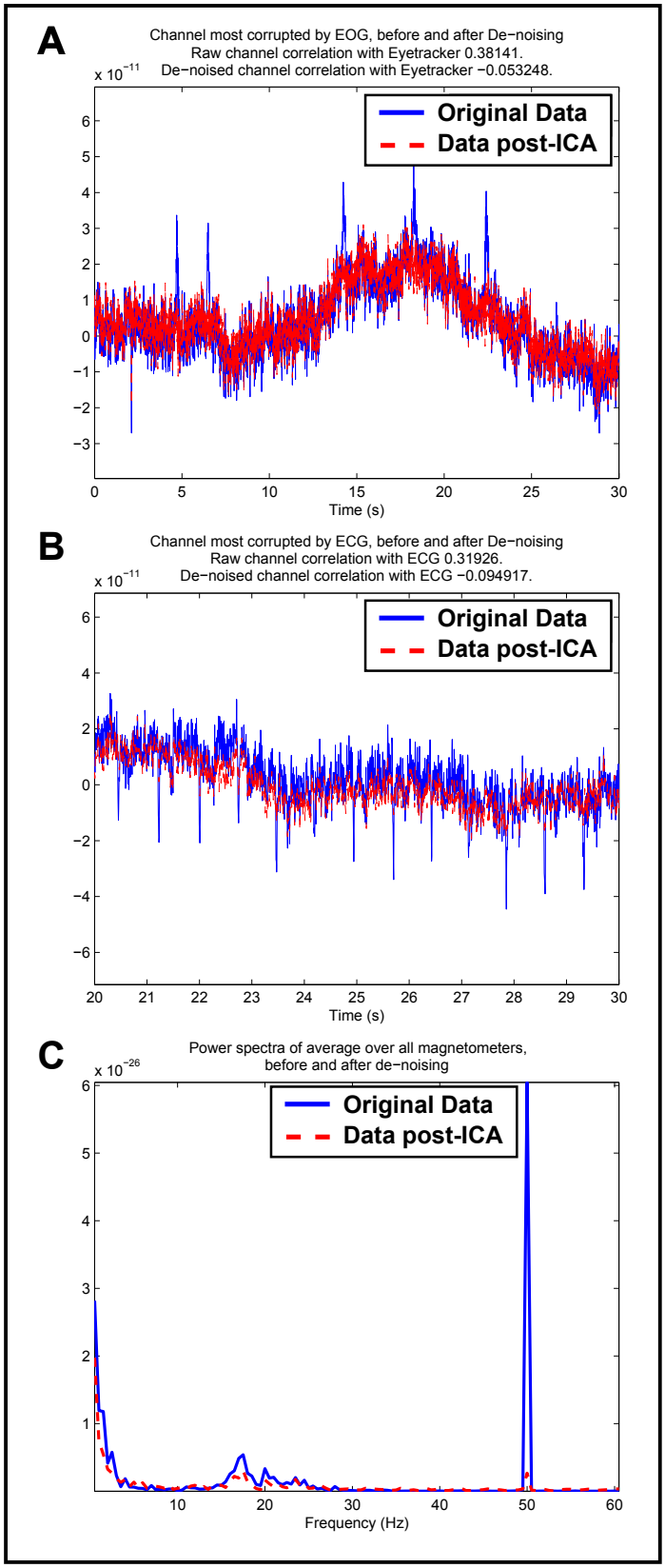


Figure 3.4: *Sensor-space demonstration of the successful removal of the eye-blink, cardiac and 50Hz mains artefacts using independent component analysis. A. 30s of data from the channel with the greatest correlation with the blink time course, measured using an optical eye-tracker, before (blue) and after (red) ICA de-noising. B. 10s of data from the channel with the strongest correlation with the ECG before (blue) and after (red) ICA de-noising. C. The power spectrum of the mean of the magnetometers before (blue) and after (red) ICA de-noising.*

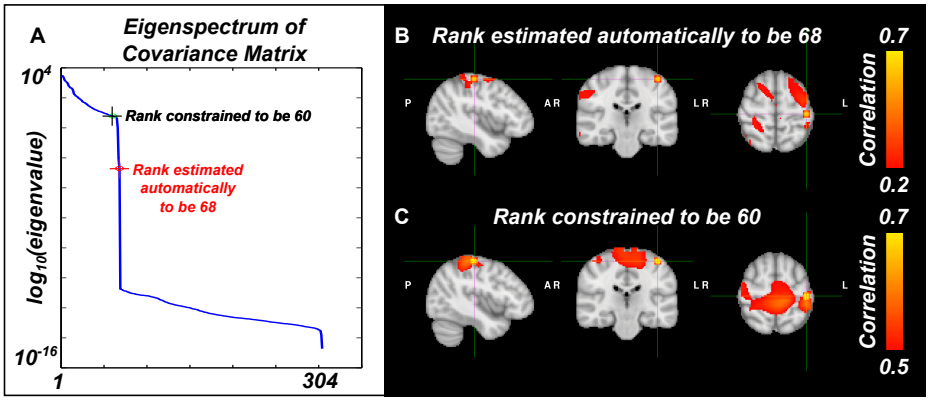


Figure 3.5: *A demonstration of the importance of correctly estimating the rank of the inverse covariance matrix when beamforming MaxFiltered data. A. The eigenspectrum of a single session of 5 minutes of resting-state data acquired with an Elekta Neuromag and preprocessed with MaxFilter^{TM}. The red cross-hair corresponds to the rank of the data estimated by the Matlab pseudo-inverse. The black cross-hair corresponds to the rank imposed by us in our modified pseudo-inverse. To demonstrate the effect of using an overestimated rank versus a correct estimate, we performed a seed-based correlation analysis between the orthogonalised, low-pass filtered envelopes at each voxel. The seed was placed in the left motor cortex (MNI coordinates $[42, -24, 60]mm$). B. The resulting correlation map using the overestimated rank of 68. No strong correlations have been found between the left motor cortex and the contralateral motor areas. C. The resulting correlation map using the user-specified rank of 60. We have found the strongest correlations with the left motor cortex to be in the medial and right motor cortices.*

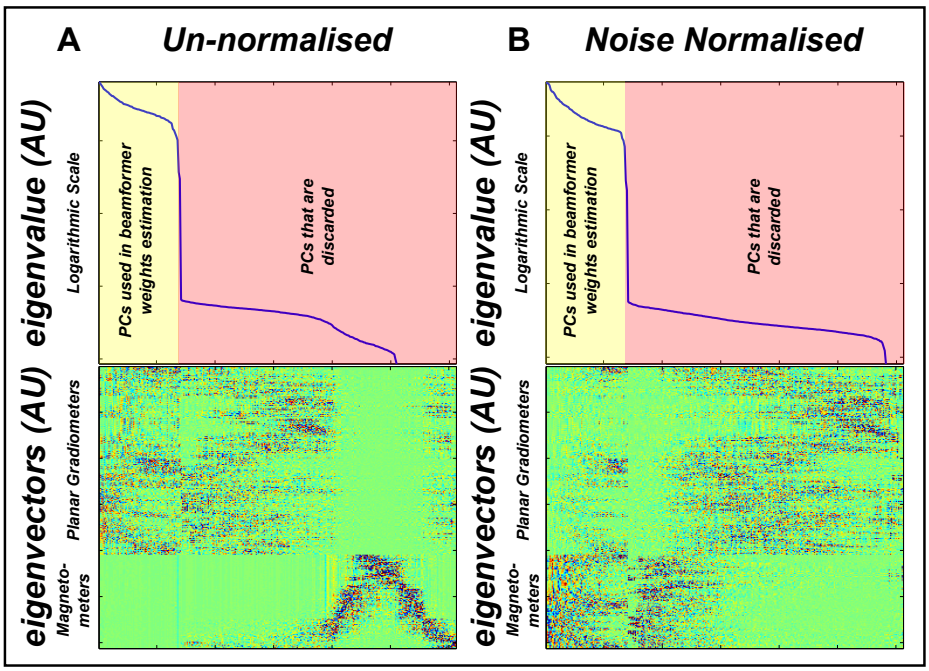


Figure 3.6: *Demonstration of the importance of normalising sensor types when beamforming MaxFiltered data. A. The eigenspectrum and corresponding eigenvectors for the unnormalised MaxFiltered data. The 64 largest principle components (PCs) contain minimal information from the magnetometers. B. The eigenspectrum and corresponding eigenvectors for the noise-normalised data (using minimum-eigenvalue sensor normalisation). Now the 64 largest PCs that will be used in the beamformer contain information from both magnetometers and gradiometers.*

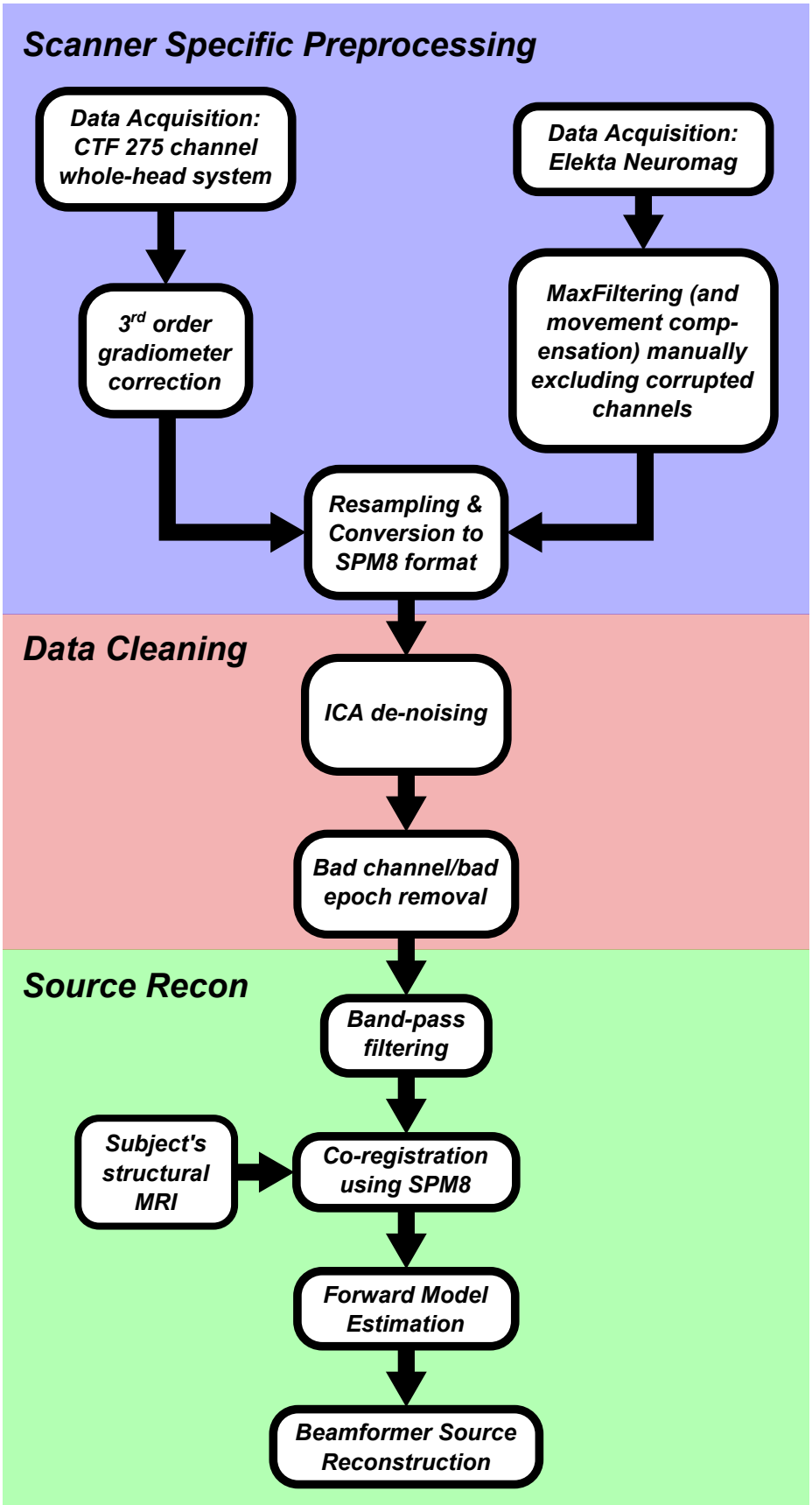


Figure 3.7: *A schematic of our preprocessing pipeline which includes far-field interference rejection, physiological artefact removal using ICA, and source reconstruction with a beamformer.*

Chapter 4

Imaging functionally-connected networks in MEG

4.1 Chapter abstract

We present three sets of functional connectivity analyses for resting-state MEG. FC is defined as correlations in the slow *co-variation* oscillations of band-limited *carrier* time series of source-space neural activity (that have been estimated with the preprocessing pipeline presented in chapter 3).

We initially investigate the optimum *carrier* and *co-variation* frequency window for measuring FC in the presence of signal leakage. We estimate the FC between 10 pairs of voxels from known RSNs to build up a distribution of FC scores as a function of *carrier*/*co-variation* frequency and compare this to an equivalent null distribution of FC scores, estimated using Monte Carlo simulations.

We proceed to develop a seed-based correlation analysis for imaging individual RSNs. We contrast several approaches for accounting for spurious FC introduced due to signal leakage. We demonstrate our seed-based analysis by imaging the sensorimotor network.

Finally, we present a framework for using independent component analysis (ICA) to decompose group-concatenated MEG envelope time courses into multiple RSNs. We demonstrate that this analysis works on MEG data acquired using a *CTF* system

and using an *Elekta Neuromag* system (which represents the first successful extraction of RSNs from *Elekta* data).

4.2 Measuring functional connectivity in band-limited power time courses

In the previous chapter, we outlined a framework for preprocessing resting-state or task-positive MEG data from its raw form through to a de-noised, source-space representation consisting of several thousand voxel time courses of neural activity for a specific frequency band. In this chapter, we will consider two types of analysis for imaging functionally-connected networks from source-space MEG data: **seed-based correlation** and **independent component analysis**.

In fMRI, the simplest technique for measuring functional connectivity between a pair of spatially separated regions is to estimate the correlation coefficient between the BOLD time courses of that pair of regions. This assumes that there is a stationary, zero-lag dependency between the two regions of interest. Due to the HRF blurring inherent to BOLD time series, stationary, zero-lag dependencies are consistently present between functionally-connected voxels. Despite its simplicity, seed-based correlation gives informative findings about functionally-connected networks [21].

However, in MEG we do not observe stationary, zero-lag statistical dependencies in the raw data and so cannot use correlation to measure FC between raw source-reconstructed time courses (because there is no blurring due to the HRF as observed in fMRI). The inability to measure FC in raw time courses of source-space MEG data with correlations could be explained by stationary, non-zero phase lags. However, coherence, a measure of FC that can account for non-zero phase lags between signals, also fails to detect connectivity in raw, source-space MEG data. This suggests that there is a high degree of temporal variability in the phase lags between voxels destroying any raw correlation/coherence. In fact, we depend on this phenomenon when beamforming as we cannot use a beamformer to source-reconstruct data containing strong temporal correlations (refer to section 2.5.3.1).

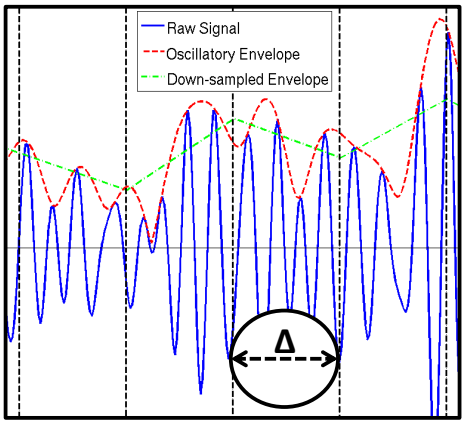


Figure 4.1: *A schematic that shows the relationship between the raw signal at a voxel obtained from beamformed MEG data (blue), its oscillatory amplitude envelope (red) and the window-averaged envelope (green) using a window length* Δ.

Instead, correlations in the band-limited power (BLP) fluctuations of a given frequency band have been shown to be an excellent metric of FC in MEG, with strong correspondence to FC measured using correlations in BOLD time courses [25]. Throughout this thesis, correlations between the slowly-fluctuating envelopes of source-reconstructed time courses are used as a surrogate measure for BLP correlations, where the envelope time course is proportional to the positive square root of the BLP time course. Figure 4.1 shows a simple example of the slow fluctuations (green) in the envelope (red) of the beamformed time course (blue). In this case, the slow fluctuations have been extracted using non-overlapping windowed averaging [25, 88]. This is very similar to a low pass-filter, with longer window lengths, Δ, corresponding to lower cut-off frequencies.

4.3 Detecting the optimum time scale to measure functional connectivity[1]

Measuring correlations between the envelopes of source-reconstructed oscillations is a robust metric of FC in MEG, as shown by Brookes et al. (2011) and Hipp et al. (2012) [25, 74]. In both these studies the authors find preliminary evidence that the

[1]The analysis in this section has been published in full by Luckhoo et al. in 2012 [88]. The descriptions of the data, methods and results have been taken from this publication.

low frequency envelope oscillations (<1Hz) are the dominant contributors to genuine FC. Brookes et al. used a windowed-average (using a window length, Δ) to extract the low frequency component whilst Hipp et al. used a low-pass filter (with a cut-off frequency, f_c. Both techniques yield equivalent results (where $\Delta \approx \frac{1}{2f_c}$). Both studies find peak FC in the beta band. We term the frequency band of the source-reconstructed neural activity the *carrier frequency* band. We term the frequency band of the envelope the *co-variation frequency* (using the terminology from [74]). The carrier frequency is often specified at the stage of source reconstruction (see Figure 3.7).

We attempted to identify what the optimum *carrier frequency/co-variation frequency* window is in which genuine FC can be measured with maximum confidence. To do this, we estimated the distribution of FC scores between 10 edges (pairs of nodes) of known RSNs as a function of *carrier frequency* band and window length, Δ. We estimated an equivalent null distribution where no genuine FC was present using Monte Carlo simulations and contrasted the two distributions using non-parametric statistics.

4.3.1 Data

Twelve subjects participated in a *2-back* working memory MEG experiment. Each subject was scanned at rest (300s during which the subject kept their eyes open and fixated on a cross) and also during a 12 minute *2-back* working memory paradigm. The data in this cohort was also used in the analysis of task-positive functionally-connected networks presented in chapter 6. Refer to section 6.3.1 for a full description of the data set.

4.3.2 Methods

For each subject in the *2-back* cohort, the beamformer was used to estimate the activity at voxels in the default mode network (DMN), right and left frontoparietal networks (FPN), sensorimotor network and hippocampi over a range of frequency

bands, between 0 and 150Hz in non-overlapping 4Hz windows. The coordinates of these nodes are listed in Table 4.1.

Network	MNI Coordinates (mm)		
Right FPN	50	-62	26
	26	32	26
Left FPN	-50	-62	26
	-26	32	26
DMN	42	-62	36
	-42	-62	36
	0	-52	28
	0	60	4
Bilateral Motor Network	41	-25	49
	-41	-25	49
Hippocampi	30	-14	-16
	-30	-14	-16

Table 4.1: *MNI coordinates of 12 nodes from five RSNs which are known to be functionally-connected. RSNs used were the right and left frontoparietal networks (FPN), the default-mode network (DMN), the motor network and the hippocampi. Table taken from [88, Table 1]*

We estimated the Pearson correlation coefficient of the windowed-average of the envelopes between the nodes within each network for a range of window lengths and for every frequency band. This gave a distribution of genuine FC scores (10 values for each subject for each frequency band/window-length). We used 13 different down-sampling window lengths ($\Delta = 0, 0.1, 0.25, 0.5, 0.7, 1, 1.5, 2, 3, 4, 5, 7, 10$s).

We also estimated the distribution of artificial null FC scores using Monte Carlo simulations. For each subject, we projected band-limited Gaussian noise through the frequency specific beamformer weights used to estimate the genuine FC scores. As there was no underlying connectivity, any measured FC was artificially introduced by cross-talk between the beamformer weights. We performed 50 instantiations per subject/frequency band. We applied the un-normalised Fisher transformation (equation 4.1) to the absolute of every correlation value. The Fisher transformation prevented any artefacts due to the unity upper bound of the Pearson correlation coefficient distorting our analysis. We took the absolute at this stage as both positive and negative correlations would be indicative of interesting FC.

$$z = \frac{1}{2} \ln \frac{1+r}{1-r} \tag{4.1}$$

In order to assess the optimum frequency band/window length Δ for measuring FC, we modelled the populations of genuine and simulated distributions as separate 2-parameter gamma distributions. We used the gamma distribution as it shares the key features of our genuine and simulated FC populations; i.e. a fixed minimum at zero, with the majority of FC scores having low values, and with a tail of strong FC scores extending to higher values.

Figures 4.2A,B,C show the histograms and gamma distribution fits for the three different analyses we performed. We performed this optimisation analysis three times: first on the resting-state block; second on the whole 12-minute *2-back* experiment (which included six 30s rest blocks and six 30s task blocks); third on the six 30s task blocks only. Qualitatively, we observed that the gamma distribution provided a sensible fit to our data in all three cases. The optimum frequency band/window length was defined as the one that maximally separated the two gamma distributions. As a metric of separability, we estimated the Fisher-transformed absolute correlation that corresponded to the 95% threshold using the gamma distribution fit to the simulated correlation values. If this value were to be used as a threshold to classify genuine from simulated FC, then one would expect a 5% false positive rate. Using this false positive rate, we then estimated the probability of correctly classifying genuine FC. This true positive rate was used as the metric of separability of the genuine FC from the spurious simulated FC: i.e. **the optimum frequency band and window length was the one that maximised the true positive rate, given a false positive rate of 5%**.

4.3.3 Results and discussion

Figure 4.2 shows the results of our investigation into the optimum *carrier frequency* and window length (proportional to the reciprocal of the *co-variation frequency*) to detect functional connectivity via correlations in envelope time courses. Figures

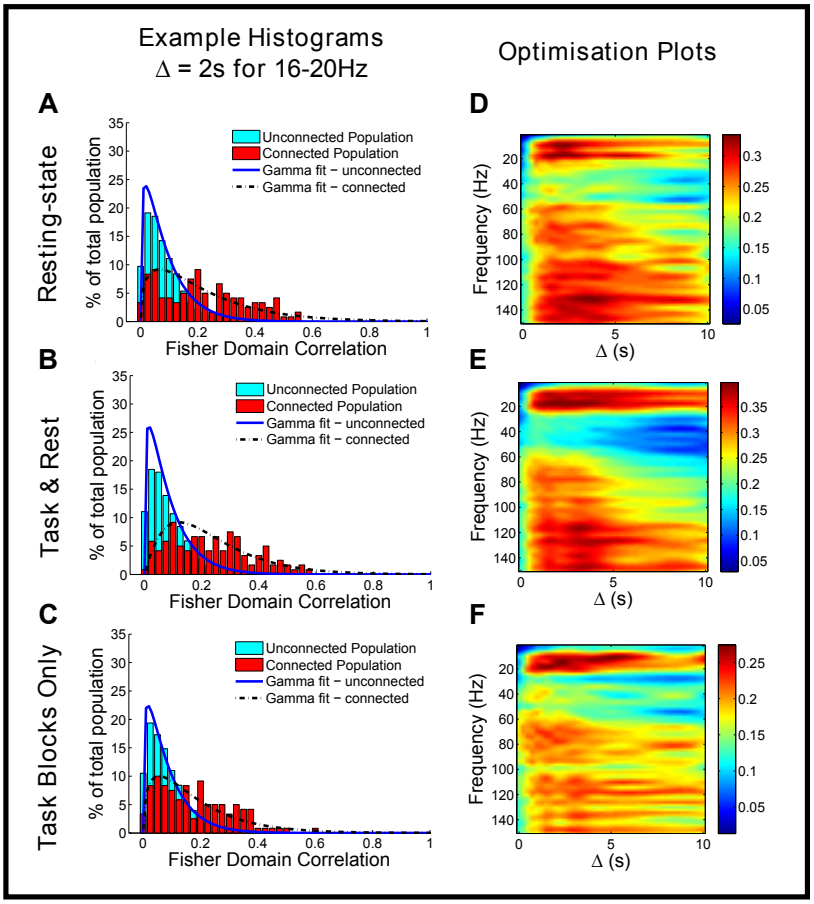


Figure 4.2: *Results of the analysis to identify the optimum down-sampling window length, Δ. The analysis has been performed on the resting-state data (A,D), the 2-back task and rest blocks (B,E) and the 2-back task blocks only (C,F). (A,B,C). Shown in red are the normalised histograms of the Fisher transformed pair-wise correlation values for 10 voxel pairs (known to be functionally-connected, listed in Table 4.1) across 12 subjects. For each set of beamformer weights, 50 unconnected time courses were simulated, giving an equivalent unconnected population whose histogram is shown in cyan. (D,E,F). The probability of true positive detection of functional connectivity (estimated at the false positive probability of 5%) for a range of frequency bands and down-sampling windows. Figure and caption taken from [88, Fig. 3]*

4.2A,B,C show that modelling the histograms of Fisher transformed, absolute correlation scores as gamma distributions was a reasonable approach that captured the critical features: 1.) a strict lower bound of zero; 2.) the bulk of the FC scores having low values; 3.) a minority of pair-wise edges that have high FC scores. Figures 4.2D,E,F show our measure of separation between the genuine functional connectivity distributions and the simulated (unconnected) null distributions. Across resting-state, task and rest blocks and pure task blocks, we found that there is an optimum frequency band to measure FC in the 8-20Hz range, in agreement with other electrophysiology studies [25, 74, 86, 91]. Lower frequency bands (<8Hz) and higher frequency bands

(>30Hz but <60Hz) both show a decreased separation between the genuine and simulated populations. Above 60Hz, we see what appears to be an increased ability to detect functional connectivity. However, this is likely to be strongly contaminated by transient muscle artefacts which we have observed to increase apparent FC measures and which are not consistently removed by our preprocessing pipeline. Our Monte Carlo simulations account for spurious connectivity introduced by the beamformer (and are equally applicable to alternative source reconstruction techniques such as minimum-norm estimates [67]). However, they do not account for physiological and environmental sources of spurious functional connectivity such as muscle artefacts, cardiac artefacts and mains interference.

We also note that there is a clear optimum range of down-sampling windows, $1s<\Delta<4s$, that maximises our rate of correctly classifying genuine FC. Again this trend is apparent in both the resting-state and task-positive state. This analysis does not explicitly investigate the presence or strength of functional connectivity, but attempts to quantify the confidence that the FC score measured between two voxels (for a given frequency band and window length) is not the product of spurious correlations introduced during source reconstruction. However, the analysis has implicitly identified a key feature of genuine FC. Spurious correlations introduced during source reconstruction will be equally present at all *co-variation frequencies*. We have identified a *co-variation-frequency*-dependent effect which must originate from the genuine functional connectivity in our data. This "sweet spot" of slow oscillations must represent a maximal point of signal-to-noise due to optimal down-sampling/low-pass filtering. As such, 1-4s down-sampling must be focusing on genuine neuronal events that drive FC occurring at that time scale. This time scale is in agreement with the time scale of FC determined by the maximally correlated window approach [43], where there is evidence of non-stationary FC occurring within 10s time windows.

Our finding is also in strong agreement with previous studies, which all indicate that the low frequency part of BLP oscillations is the best for measuring functional connectivity over several minutes [25, 74, 86]. Brookes et al. showed a similar result comparing windowed-average correlations between voxels in the sensorimotor network

[25]. However, their analysis only considered the sensorimotor network and did not quantify the separation between genuine and spurious FC as we have.

Hipp et al. performed a complementary analysis where they measured resting-state functional connectivity in the sensorimotor, auditory and visual networks [74]. They accounted for spurious functional connectivity by orthogonalising pairs of voxels prior to estimating the envelopes (refer to section 4.4.1 for a full description of orthogonalisation) [29, 74]. Furthermore, instead of using a windowed-average, they applied a low-pass filter to focus on the slow fluctuations. They found that the 8-32Hz range was optimum for measuring FC, in strong agreement with our finding. They also found that FC dropped off with low-pass cut-off frequencies above 0.32Hz (which approximately corresponds to using $\Delta < 1$s). This drop off in FC with lower cut-off frequency corresponds exactly to our finding that FC drops off with $\Delta < 1$s.

However, Hipp et al. found that FC scores increased with more severe low-pass filtering (up until 0.032Hz (equivalent to $\Delta = 15$s which is the limit of their analysis). It is likely that this is because their analysis failed to account for the decrease of the number of observations when greater low-pass filtering is applied. Under the null hypothesis, the standard error of correlation values is proportional to $\frac{1}{\sqrt{N}}$ where N is the number of independent observations. As we are considering the absolute of the correlation coefficient, the mean correlation under the null hypothesis is also proportional to $\frac{1}{\sqrt{N}}$. We have implicitly accounted for this effect with our Monte Carlo simulations (as have Brookes et al.). However, Hipp et al. did not use Monte Carlo simulations and, as such, only observed an ongoing increase in FC with more low-pass filtering.

From this analysis, we concluded that the ideal down-sampling window length is approximately 2s, equivalent to a *co-variation frequency* of $<$0.5Hz. In the ICA framework described in section 4.5.3 we use the non-overlapping windowed-average to estimate the low frequency envelope time courses. Accordingly, we use $\Delta = 1$s or $\Delta = 2$s in all our resting-state ICAs. In chapter 6, we use $\Delta = 0.5$s when analysing a *2-back* task with 2s long trials. This is because we wanted to retain within-trial temporal resolution. When using seed-based correlation analyses (described in section

4.4.2), we use a low-pass filter with a cut-off frequency of 0.5Hz to focus on the optimum low frequency *co-variation frequencies*.

4.4 Measuring FC with seed-based correlation

Seed-based correlation has been widely used in MEG functional connectivity analysis to measure both stationary and non-stationary FC [25, 43, 44, 74]. Conceptually, the analysis is extremely simple: one selects a seed voxel (or region) and extracts a summary envelope time course. One then estimates the correlation between that envelope time course and the envelope of every other voxel to obtain a correlation map specific to that reference seed.

4.4.1 Accounting for correlations due to signal leakage

One challenge with seed-based connectivity as a technique is that the resulting correlation maps contain strong artificial contributions caused by spurious correlations due to signal leakage between source-reconstructed voxels (refer to section 2.6 for a full description of signal leakage). It is essential to account for these artificial correlations. If one ignores them, then one can falsely detect FC even when none is present. We consider three techniques for accounting for spurious correlations.

Spatial separation: The first is very simple. We assume that spurious correlations are predominantly introduced between neighbouring regions. This technique is used by de Pasquale et al. [43]. They stated that correlations measured between the left ventral and posterior intraparietal sulcus (separated by 19mm) were not due to signal leakage because there was a minimum in the profile of correlation scores along the line connecting the two nodes. Furthermore, they noted that other node pairs were sufficiently separated to be within the spatial resolution of MEG. This line of reasoning assumes that spurious correlations drop off in a well-behaved, isotropic fashion with increasing distance from the seed.

However, such behaviour is not always observed in MEG. Here, we demonstrate

that artificial correlations can exist over long distances, and with an anisotropic spatial profile by using correlations between beamformer weights as an approximate measure for cross-talk induced correlations [65]. Figure 4.3A shows the correlation between the beamformer weights at a seed in the posterior cingulate cortex (PCC) and the beamformer weights at every other voxel for a single session of eyes-closed resting-state data, beamformed in the 13-30Hz band (10 minutes acquired on a *CTF* 275-channel whole-head system - for a full description of the data, refer to 5.5.1.1). We can see that the weights correlation does not continuously fall off with distance in an isotropic fashion from the seed. Figure 4.3B shows the correlation between the low-pass filtered ($<$0.5Hz) envelopes of the same 13-30Hz data with the seed voxel. The correlation pattern follows the profile of the weights correlation. These two results show that it is dangerous to assume how much signal leakage is present simply from the distance between the two voxels and the presence of a minimum in the correlation scores in the voxels directly separating them. We conclude that distance from the seed is a misleading technique for accounting for the contribution of artificial correlations to a given FC measurement, and hence do not use it.

Monte Carlo simulations: The second technique for accounting for spurious functional connectivity is to use Monte Carlo simulations. Monte Carlo simulations have been successfully used in multiple studies and we use such simulations in section 4.3 [25, 30, 88]. However, they are computationally expensive and therefore are typically limited to investigating the contribution of spurious FC for a small number of edges (pairs of regions) and do not scale well to whole-brain seed-based correlation analyses.

Orthogonalisation: The final technique is known as *orthogonalisation* or *zero-lag correction* [29, 74]. Consider a pair of voxels: before estimating the envelope time course for each voxel we can orthogonalise one voxel with respect to the other. Correlations caused by signal leakage are zero-lag (or close to it) and therefore are eliminated. We can then estimate correlations between the envelope time courses of the

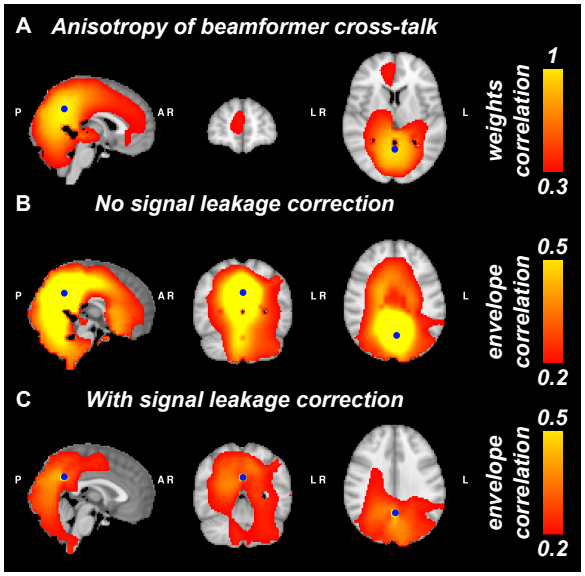


Figure 4.3: *A single session demonstration of the anisotropic spatial profile of correlations due to signal leakage. 10 minutes of eyes-closed resting-state data was beamformed in the β-band (13-30Hz). A. The correlation between the beamformer weights at a seed in the posterior cingulate cortex (PCC) (MNI: $[0, -54, 28]mm$) and every other voxel. B. The correlations between the low-pass filtered ($<0.5Hz$) envelopes using the same seed without accounting for correlations due to signal leakage. Note that the weights correlation and envelope correlation maps are extremely similar. C. The same correlation map except accounting for signal leakage by orthogonalising each seed/target voxel pair before estimating the envelope. Note that the FC between the PCC and the anterior cingulate cortex (ACC) has been removed, indicating that it was predominantly due to signal leakage.*

orthogonalised time series with confidence that signal leakage has been accounted for. However, this technique will remove any genuine zero-lag effects and so is potentially over-conservative. Orthogonalisation is substantially more efficient than Monte Carlo simulations and so is well suited to whole-brain seed-based correlation analyses. One point to note is that the orthogonalisation of two voxels is asymmetric: the envelope correlation between voxels a and b after a has been orthogonalised with respect to b is not equal to the envelope correlation after b has been orthogonalised with respect to a. The simplest way to account for this asymmetry is to estimate both correlations and take the mean [74]. Figure 4.3C shows the equivalent orthogonalised correlation map for the same single session of 13-30Hz data. The orthogonalisation has removed a substantial amount of apparent FC from near the seed and furthermore has removed the apparent FC between the PCC and the anterior cingulate cortex (ACC) that is present in the standard correlation map (when using the same correlation thresholds).

For all the seed-based correlation analyses in this thesis we always apply orthogo-

nalisation in order to eliminate spurious correlation due to signal leakage (excluding the optimisation analysis in section 4.3 where we used Monte Carlo simulations).

4.4.2 Seed-based correlation pipeline

In this section, we present the analysis pipeline stages of our seed-based correlation analyses. This analysis follows on directly from the preprocessing pipeline described in section 3.6. Starting with a band-limited time course of neural activity for each vertex of a 6mm grid spanning the whole brain, we specify a seed voxel which is fixed across the group. We then perform the following stages (summarised in Figure 4.4) at every voxel:

1. **Orthogonalisation**: We orthogonalise the seed and target voxels via a single regression [29, 74]. We then perform stages 2-4 for both the *seed-orthogonalised-to-target* and *target-orthogonalised-to-seed* time courses.

2. **Enveloping**: We estimate the oscillatory amplitude envelope for the orthogonalised time courses by computing the absolute of the *analytic signal* computed using the Hilbert transform [92].

3. **Low-pass filtering**: We low-pass filter the data to remove *co-variation frequencies* over 0.5Hz.

4. **Correlation**: We estimate the Pearson correlation coefficient between the low-pass filtered envelopes.

5. **Averaging**: We average the two correlation scores (for the *seed-orthogonalised-to-target* and *target-orthogonalised-to-seed* pairs.

This gives a correlation map for every subject/session of data. We take a simple average over subjects/sessions to get a group correlation map. In principle this framework can easily be implemented in a mixed-effects general linear model [139]. This would properly model uncertainty at the session, subject and group levels to allow proper statistical inference. However, in this thesis we develop alternative techniques for performing group statistics on functionally-connected networks (described

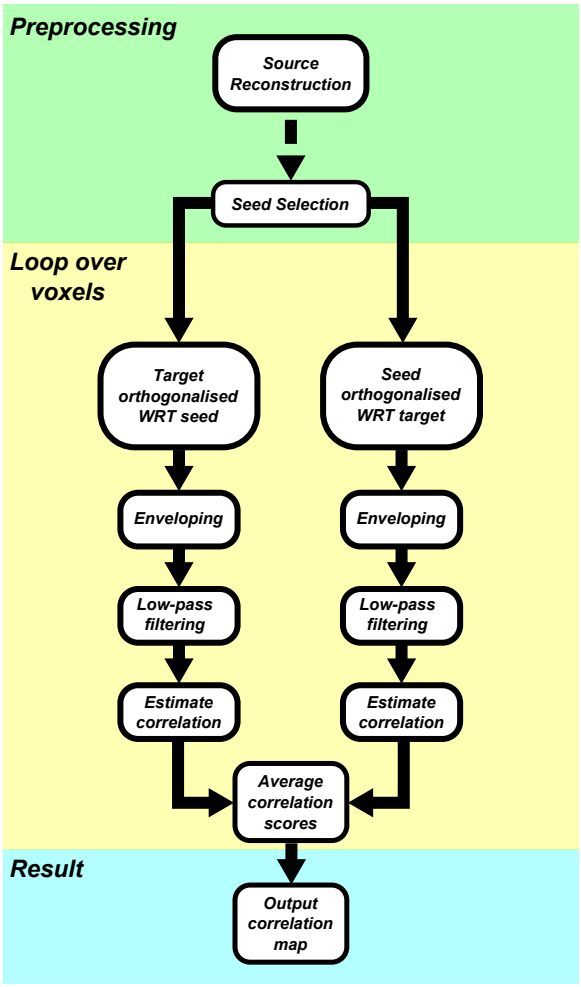


Figure 4.4: *A schematic of the seed-based correlation analysis used in this thesis. Note that for each seed-target voxel pair, we have to perform two orthogonalisations and estimate two subsequent correlations in order to account for the asymmetry of pairwise FC measured between orthogonalised BLP time courses. WRT: with respect to.*

in chapters 5 and 6). As such, simple group mean correlation maps are sufficient for validating our preprocessing and source reconstruction stages.

4.4.3 Results

Imaging the sensorimotor network: To demonstrate the efficacy of our seed-based correlation analysis, we applied it to nine 10-minute sessions (from nine subjects) of *eyes-open* resting-state data, acquired on a *CTF* system. For full details of the data acquisition, refer to section 5.5.1.1. Each subject was preprocessed and source-reconstructed using the pipeline described in section 3.6 to give estimates of the time courses of β-band neural activity over a 6mm grid. We applied the seed-correlation analysis described in section 4.4.2 using a voxel in the left motor cortex

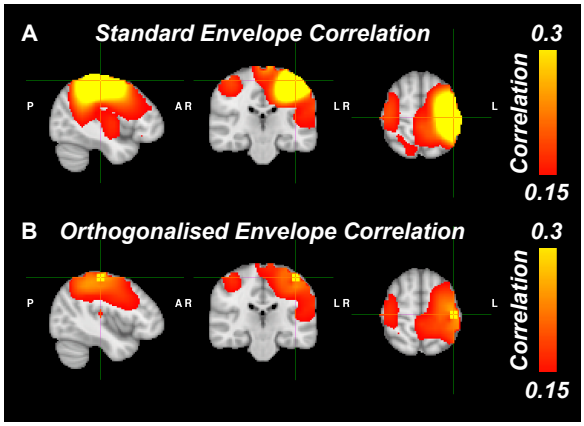


Figure 4.5: *Group average correlation maps from a seed-based correlation analysis performed on 9 subjects. For each subject the correlation between the low-pass filtered (<0.5Hz) envelopes of the β-band neural activity at each voxel and a seed voxel in the left motor cortex (MNI:* $[-42, -24, 60]$ *mm) was estimated. A. The resulting correlation map without any correction for signal leakage. B. The resulting correlation map using orthogonalisation to remove any correlations due to zero-lag signal leakage.*

(MNI: $[-42, -24, 60]$ mm) as the seed. We estimated the correlation maps without, and with, the orthogonalisation correction for signal leakage. Figure 4.5 shows the group average correlation maps without orthogonalisation (A) and with orthogonalisation (B). We have reproduced the well-documented functional connectivity between the left and right motor areas which we know to be detectable in the slow envelope fluctuations of the β-band [25]. Secondly, we have demonstrated how much signal leakage contributes to the apparent functional connectivity. In Figure 4.5A the correlation map is dominated by the strong correlations near the seed voxel. These correlations are undoubtedly due to signal leakage. After orthogonalisation, we find the local correlations near the seed to be reduced but not completely removed, indicating that there is some non-zero-lag statistical dependencies indicative of FC; this is in exact accordance with previous demonstrations of orthogonalisation correction [29, 74]. Most importantly, this result validates our preprocessing pipeline, in particular confirming that our beamformer implementation was performing correctly.

Seed-selection sensitivity: We performed an additional seed-based correlation analysis on a single session of eyes-open resting-state data to investigate the sensitivity of seed-based analyses to seed selection. The session was taken from the APOE cohort (described in section 5.5.1.2). It consists of six minutes of eyes-open resting-state

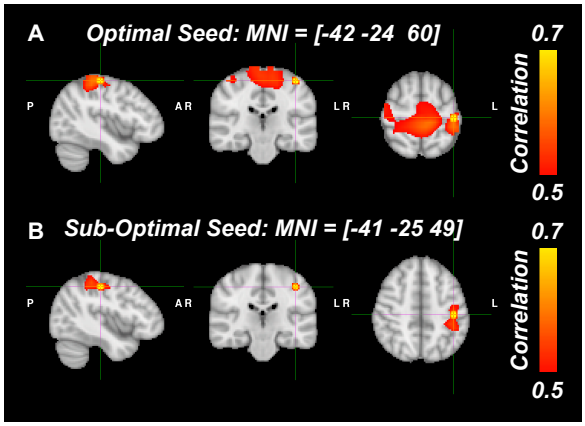


Figure 4.6: *Correlation maps from a seed-based correlation analysis of a single subject of eyes-open resting-state data (6 minutes acquired on an Elekta Neuromag). Two seeds were placed in the left motor cortex and the correlations between the low-pass filtered (<0.5Hz) envelopes of the orthogonalised β-band oscillations were estimated between each seed and the rest of the brain. A. The correlation map when using the optimum seed for that location (MNI: $[-42, -24, 60]mm$). B. The correlation map when a sub-optimal seed is used (MNI: $[-42, -25, 49]mm$ taken from [25]).*

data, acquired on an *Elekta Neuromag*. We seeded in the left motor cortex (MNI: $[-42, -24, 60]$mm) and performed the same seed-based correlation analysis, described above (this analysis is identical to the seed-based correlation result in Figure 3.5). We selected this seed voxel by picking the peak voxel in the fMRI-derived spatial ICA map of the sensorimotor network (using the spatial ICA results from [122]). We then used a different seed (MNI: $[-41, -25, 49]$mm), taken from a group-average MEG-motor localiser task (reported in [25]) and repeated the analysis. Figure 4.6 shows the results of these almost identical analyses. In Figure 4.6A, we successfully detected functional connectivity in the right motor cortex, agreeing with previous studies [25, 74]. However, when we used the second seed and applied an otherwise identical analysis (including the correlation thresholds used when visualising the maps), we failed to detect any cross-hemisphere FC. This "sub-optimal" seed was taken from Brookes et al.'s 2011 MEG investigation of functional connectivity between the left and right motor areas, emphasising that using seeds from previous studies does not guarantee good quality correlation maps [25].

4.4.4 Summary: strengths and weaknesses of seed-based correlation

The major weakness of seed-based correlation approaches are their sensitivity to the seed selection. The choice of seed can greatly modify resulting measures of functionally-connectivity. Furthermore, when comparing two correlation maps with different seeds, determining which map is a better reflection of the underlying functional connectivity is far from trivial. Seed selection notwithstanding, seed-based correlation is an extremely useful tool for investigating functional connectivity in MEG. It is simple and fully deterministic (unlike ICA which often uses a random initialisation and so can give different results when repeating the analysis). In this thesis, we have used seed-based correlation extensively as a tool for validating our analysis pipeline.

4.5 Imaging functional networks with ICA

4.5.1 Independent component analysis

In section 1.5.1 we introduced independent component analysis (ICA) as a blind source separation technique capable of extracting multiple RSNs from resting-state fMRI data [13]. Here, we go into more detail on what ICA is and how it works.

4.5.1.1 The linear mixing problem

ICA can be thought of as a linear un-mixing algorithm. Given a set of I source signals each with T samples, we can express these sources as an $I \times T$ matrix $\mathbf{S}$. We define the matrix of mixed signals, $\mathbf{X}$ (which is Q mixed signals $\times$ T samples and where $Q \geq I$), as a linear mixture of the sources $\mathbf{S}$. Finally, we define the mixing matrix, $\mathbf{A}$, which describes how the source signals are mixed together to produce the signal mixture. Equation 4.2 describes the linear mixing of $\mathbf{S}$ to produce $\mathbf{X}$:

$$\underbrace{\mathbf{X}}_{\text{signal mixture}} = \underbrace{\mathbf{A}}_{\text{mixing matrix}} \times \underbrace{\mathbf{S}}_{\text{sources}} + \underbrace{\boldsymbol{\varepsilon}}_{\text{noise}} \tag{4.2}$$

where $\boldsymbol{\varepsilon}$ is a Q signals $\times$ T samples matrix of additional noise. Let us assume that we know $\mathbf{X}$: perhaps we have physically measured some signal mixture in an experiment. We would like to estimate the set of un-mixed sources, $\mathbf{S}$. If we knew the mixing matrix, $\mathbf{A}$, then estimating $\mathbf{S}$ would be trivial. However, if both $\mathbf{A}$ and $\mathbf{S}$ are unknown then the problem appears intractable.

4.5.1.2 Independence and the central limit theorem

ICA solves this apparently intractable problem by assuming that the I sources are all statistically independent. Put simply, the *central limit theorem* tells us that a linear mixture of statistically independent signals will have a more Gaussian distribution than any of the underlying sources. ICA essentially reverses this idea: given a linear mixture of I statistically independent sources, **the set of time series that each have maximal *non-Gaussianity* must correspond to the original underlying sources** [81].

From this statement, we can see how an ICA algorithm might be formulated. Given our mixture $\mathbf{X}$, we can guess an initial mixing matrix and estimate the corresponding set of I sources. We can then iteratively modify our mixing matrix and search for the set of maximally non-Gaussian components. Two obvious questions arise: 1.) How do we measure non-Gaussianity? 2.) How do we know how many components to seek?

The simplest measure of non-Gaussianity is kurtosis [79]. However, different implementations of ICA use different measures in order to improve accuracy, robustness or speed of convergence. Correctly estimating the number of sources is not trivial and we discuss this in section 4.5.1.4.

4.5.1.3 Implementing ICA

***Preprocessing*:** In our initial formulation of the linear mixing model (equation 4.2), we imposed a constraint that the number of mixed signals is greater than or equal to the number of mixed signals: $Q \geq I$. Many ICA implementations actually force the constraint that number of signals equals the number of sources: $Q = I$. This makes the resulting reduced mixing matrix, $\mathbf{A_{square}}$, an $I \times I$ square matrix and hence is known as squaring [81]. Squaring is achieved by performing a dimensionality reduction on the signal mixture (typically using principle component analysis (PCA) to discard all but the I largest principle components). This is also assumed to remove the noise contribution $\boldsymbol{\varepsilon}$ resulting in the simplified noise free mixing model:

$$\mathbf{X_{reduced}} = \mathbf{A_{square}}\mathbf{S} \tag{4.3}$$

As well as performing a dimensionality reduction, the PCA preprocessing is commonly used to pre-whiten or sphere the surviving principle components by normalising each one to have unit variance. This dramatically simplifies the ICA decomposition for the following reason: if the signal mixture is uncorrelated and whitened then the mixing matrix that maps the independent sources onto the whitened mixture must be an orthogonal matrix [78]. This reduces the degrees of freedom of the ICA problem from I^2 to $\frac{I(I-1)}{2}$.

PCA preprocessing has certain limitations. Firstly, it does not actually provide any information on what the optimal model order selection is. Secondly, it assumes that all the source information is contained in the I largest PCs which is not guaranteed [81]. In section 3.3.1, we encountered this issue when performing PCA preprocessing on multi-sensor-type data, where all the information of the magnetometer sensors are contained in the smallest 102 PCs. We developed minimum eigenvalue normalisation specifically to counter this issue.

INFOMAX: The **INFOMAX** algorithm was developed in 1995 by Bell and Sejenowksi [16]. It essentially uses the information-theoretic measure known as *negentropy* (or differential entropy) as a metric of non-Gaussianity, combined with an information-maximisation algorithm [81].

FastICA: **FastICA** is an implementation of ICA which uses a fixed-point algorithm to estimate the ICA decomposition [77, 79]. As such, the **FastICA** algorithm has cubic convergence properties and performs substantially faster than gradient-ascent optimisation methods. Furthermore, it does not require the specification of a learning rate. This is an advantage as sub-optimal learning rates can destroy ICA convergence [78]. **FastICA** can be set up to maximise kurtosis [79] or *negentropy* [77].

There are other implementations of ICA which have not been discussed here. For an extensive review of ICA implementation strategies, refer to [78].

***Selecting an ICA implementation*:** In this thesis, we use the **FastICA** implementation of ICA, via the **FastICA** open-source *Matlab* toolbox (*Department of Information and Computer Science, Aalto University*).[2] **FastICA** has three critical design choices: **model order** (which we discuss in section 4.5.1.4), **contrast function** and **decomposition approach**.

When using **FastICA**, it is necessary to specify a non-quadratic **contrast function** (**FastICA** has a selection hard-coded in). We use the *tanh* **contrast function**, which is recommended for general purpose applications [77]. The other **contrast functions** are recommended when strongly super- or sub-Gaussian independent components are being estimated. As we often need to estimate both super- and sub-Gaussian components (such as the cardiac and mains interference components described on section 3.3), we use the most robust **contrast function** *tanh*. The **decomposition approach** refers to how an ICA algorithm estimates multiple ICs. ICA algorithms can attempt to find a single IC at a time, orthogonalising the signal

[2]The toolbox can be downloaded for free at `http://research.ics.aalto.fi/ica/fastica/`.

mixture to each one as they are found. This is known as a *deflation* approach. Alternatively, the algorithm can attempt to estimate all the ICs simultaneously, known as a *symmetric* approach. *Deflation* approaches have the weakness that errors in the first components that are estimated compound to worsen the estimates of subsequent ICs [78]. As such, we utilise the *symmetric* estimation approach in **FastICA**.

One reason for selecting **FastICA** is that it has enjoyed wide success in RSN imaging applications. The **MELODIC** ICA implementation (*FSL*, *FMRIB*, *Oxford*) uses the **FastICA** fixed-point algorithm [15]. **MELODIC** is the ICA technique used in the fMRI imaging studies that we use as a gold standard for RSN maps throughout this thesis [13, 122]. Furthermore, Mantini et al. (2011) performed a comparison of the **FastICA**, **INFOMAX** and **SOBI** ICA implementations using simulated data sets and found **FastICA** to have the best performance. [17, 90]. Finally, Brookes et al. (2011) successfully employed **FastICA** to image 8 RSNs in source-reconstructed MEG data [30]. As one aim of this thesis was to independently validate this finding, utilising the same ICA algorithm was an obvious choice.

4.5.1.4 Estimating the model order in ICA

Throughout our discussion of ICA so far, we have assumed that the number of sources, I, is known. In many applications of ICA, the true number of underlying sources (also known as the model order) is not known *a priori*. Imaging resting-state networks is no exception. It is necessary to estimate the model order prior to the ICA decomposition. Overestimating the model order can force the ICA to break up RSNs into sub-networks, while underestimating the model order can force spatially-distinct RSNs to be merged together. Here we present a selection of strategies for estimating the model order.

Using the full rank: The simplest approach is to use the full rank of the data as an estimate of the number of underlying sources. For an idealised, noise-free linear mixing process, this is a reasonable assumption.

FastICA's *deflation* approach: ICA algorithms have iterative optimisation procedures, with user-specified maximum iterations built-in to stop the program when the optimisation fails to converge. By implementing a *deflation* approach, one can keep extracting ICs until either the full rank of the data is reached, or the ICA algorithm fails to converge [89]. The issues with this approach are twofold. Firstly, failure to converge after a set number of iterations could occur for reasons other than when the number of sources has been reached. Secondly, it still requires an arbitrary definition of failed convergence (e.g. > 1000 iterations).

MELODIC's Bayesian framework for model order estimation: MELODIC attempts to estimate the number of noise sources present that inflate the rank of the signal mixture beyond the number of underlying sources [15]. By assuming that the sources are non-Gaussian but the noise contributions are Gaussian, it can approximate the true model order.

Running multiple ICA decompositions with different model orders: Perhaps the most heuristic approach is to run numerous ICAs at a range of model orders. Identifying components that consistently recur at higher and higher model orders implies that they represent genuine single underlying sources. Furthermore, one can map the break-up of components into sub-components and infer some degree of functional taxonomy from the multiple ICA outputs [1]. However, this approach is computationally intensive. Furthermore, matching components from multiple ICA runs that have been performed using different model orders is not trivial.

Using an *a priori* hypothesis: *A priori* hypothesis-driven approaches are the alternative to empirical, data-driven approaches for selecting a suitable model order. Such methods use prior beliefs or knowledge about the number of underlying sources. In MEG RSN analysis, this is the most widely adopted approach as the aim of most MEG analyses to date has been to replicate the functional networks that have already been identified in fMRI [26,30,88]. As such, it is common to select ICA model orders of 20-25 as these have repeatedly been shown to yield physiologically sensible RSN

maps in fMRI studies [13, 122].

The issue of model order selection has not been and may never be satisfactorily solved. Only in the small set of cases where the ground truth is known is model order selection not an issue. In all other cases, a combination of data-driven approaches, hypothesis-driven approaches and explicit investigations into the effect of changing model order can hope to settle this issue. Ultimately, the "best" model order may simply depend on what the ICA decomposition is going to be used for.

4.5.2 Applying ICA to source-reconstructed MEG data

In chapter 3 we applied ICA to single sessions of MEG data to extract and remove sources of physiological and environmental interference. Here, we outline exactly how we can utilise ICA to detect functionally-connected networks. As we have previously established, functional connectivity in MEG is best detected by considering the slow fluctuations in the envelopes of band limited neural activity. Consequently, when applying ICA, we feed in the windowed-average of the oscillatory amplitude envelope. Essentially, we aim to decompose these envelope time courses into a set of spatial maps and associated time courses.

4.5.2.1 Group ICA using concatenation

In fMRI, group ICA is used to extract RSNs from fMRI data acquired from multiple subjects/sessions. Group ICA is used for two reasons. Firstly, to have enough samples to perform ICA robustly typically requires around an hour of data. It is extremely challenging to acquire one hour of high quality resting-state data in a single session, although the data could be gathered over multiple sessions and then combined in a single ICA. Furthermore, in general, we are not interested in subject-specific RSNs, but rather the group average. Given enough data for each subject, one might propose performing multiple subject-specific ICAs after which the subject-specific RSNs could be averaged. The challenge with such an approach is establishing correspondence between the ICs from one ICA and another. As each decomposition is performed

in isolation, there is no guarantee that a one-to-one mapping across the group will occur.

Instead, the resting state data is temporally concatenated and a single ICA decomposition performed to get a group-average set of spatial maps and concatenated time courses: this forces correspondence in the components across the group. Figure 4.7 shows a schematic of the temporal concatenation and ICA applied to MEG data.

4.5.2.2 Group ICA using concatenation in MEG

We use group ICA when analysing MEG data for exactly the same reasons as in fMRI. However, there are two key additional considerations to make in MEG: *weights normalisation* and *de-meaning*. We found that it is necessary to apply *weights normalisation* to the beamformed MEG data prior to ICA decomposition. Without *weights normalisation*, there is a spatially varying bias in the contribution of sensor noise across the brain. As Figure 2.4 shows, without *weights normalisation* this bias in the noise estimate leads to peak variance of the neural activity and its envelope in the deep regions of the brain. When performing the PCA dimensionality reduction, these high variance, noisy voxels will dominate the largest principle components and genuine neural activity will be squeezed into the lower variance PCs and discarded. *Weights normalisation* down-weights the deeper voxels and provides an unbiased estimate of the sensor noise contribution, which has been previously shown to be a critical step in ICA when applied to fMRI data [15]. As such, the noisy deep brain regions will no longer dominate the largest PCs. Figures 4.11A and B show group ICA outputs with and without *weights normalisation*.

In fMRI, the mean of a voxel's BOLD time course does not reflect an absolute value and so is discarded. In MEG, the mean of the envelope during a session can be directly related to the mean oscillatory power of that frequency band. As such, it is tempting to include the mean of the envelope at each voxel for every session. However, as we show in section 5.3, applying session-specific *weights normalisation* has the consequence of destroying the mean of the envelope as an absolute measure of oscillatory power. As it is no longer an absolute measure, we cannot combine *weights-*

normalised means from multiple, separately beamformed sessions of data. Therefore the *weights-normalised* envelopes at every voxel for each session must be de-meaned prior to temporal concatenation. The MEG ICA will then be driven purely by the variability over time, in the same manner as fMRI. Figure 4.11C shows the effect of including the *weights-normalised* means in the ICA.

4.5.2.3 Spatial versus temporal ICA

We have established ICA as a technique for decomposing source-reconstructed MEG data into a set of spatial maps (some of which correspond to RSNs) and associated time courses (corresponding to the time course of each RSN) by finding a set of maximally independent components. We are presented with a critical choice: should independence be maximised in the spatial dimension (known as spatial ICA) or in the temporal dimension (known as temporal ICA).

In fMRI, spatial ICA is the dominant technique. This is because fMRI data has excellent spatial resolution and so the spatial dimension is extremely well sampled. It has much poorer temporal resolution. By performing spatial ICA, the number of samples is equal to the number of voxels, whilst the number of mixed signals is equal to the smaller number of BOLD time points. Each independent component corresponds to a vector of values that make up a spatial map. The column of the mixing matrix is a time course that weights the contribution of that map to each time point in the BOLD data. The spatial maps are forced to be independent which makes them (mostly) non-overlapping.[3]

In MEG, we have poor spatial dimensionality (as low as 64 in *MaxFiltered* data). MEG has much higher temporal dimensionality, even after aggressive windowed-average down-sampling. Spatial ICA generates single region spatial maps rather than multi-node networks. Therefore, temporal ICA is utilised to image whole RSNs. In chapter 7 we use spatial ICA to extract a functional parcellation made up of single network node spatial maps. Temporal ICA outputs a set of temporally independent

[3]ICA decompositions that try to maximise spatial non-Gaussianity also maximise uncorrelatedness and sparsity in the spatial maps forcing them to be approximately non-overlapping [42].

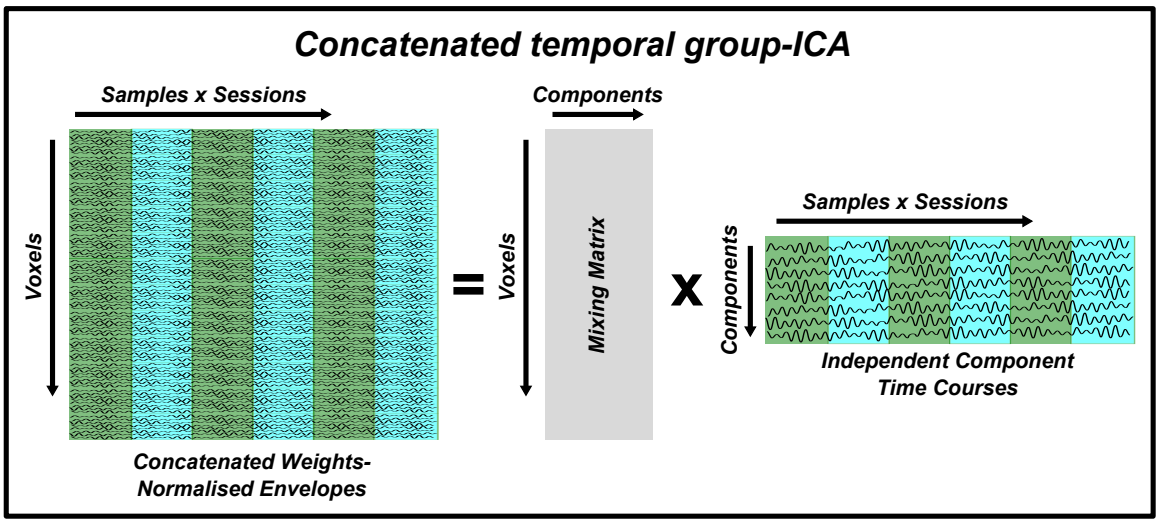


Figure 4.7: *A schematic of the decomposition of temporally concatenated envelope time courses using temporal group ICA.*

time courses and associated spatial maps (which correspond to the mixing matrix). The time courses are mutually uncorrelated but the maps can be overlapping.

4.5.2.4 Visualising resting-state networks after performing ICA

There are numerous options for visualising the RSNs that are found after performing group ICA on the window-averaged envelopes. Here, we briefly discuss three options:

Normalised covariance maps: This corresponds to visualising the columns of the mixing matrix $\mathbf{A}$. The temporally independent time courses are forced to be unit variance. Each column of the spatial map weights the contribution of that unit variance time course to the envelope time course and is equal to the covariance between the independent time course and the concatenated voxel time course (because the ICs are all uncorrelated). In principle, the raw spatial maps describe a physical relationship and have absolute units. However, as a consequence different thresholds are required for different components which introduces an additional complexity, especially when comparing across multiple modalities. To avoid this issue, we normalise each covariance map to have unit variance prior to visualisation, thus allowing the same threshold to be used across all the components. We call these "normalised covariance maps" and they are directly comparable to seed-based covariance maps [65].

Correlation maps: By estimating the Pearson correlation coefficient between each independent time course and the concatenated envelope time course, we produce a map of correlation scores for each component. The advantage of correlation maps is that they can be directly compared with correlation maps from seed-based analyses, except instead of using a seed-voxel, we use a seed component that was learnt from the data. Correlation maps have been used by several MEG ICA studies [30, 88].

Statistical maps: One technique for visualising RSN maps is to model the ICA spatial maps as a mixture of distributions [52]. One such model uses a Gaussian distribution to model inactive voxels and one (or two) Gamma distribution(s) to model the positive (and negative) active voxels [123]. This technique converts the spatial map values into statistical values. However, it still requires the specification of an arbitrary threshold.

We use the normalised covariance and correlation maps in this thesis as they are simple to implement and readily comparable to seed-based analyses. In chapters 5 and 6, the statistical frameworks that we develop utilise the raw spatial maps and time courses. Consequently, the choice of spatial map visualisation is of no consequence.

4.5.3 A framework for applying group ICA to MEG data

In this section, we outline the specific stages of our group ICA method. This analysis follows on directly from the preprocessing pipeline described in section 3.6. We perform the following stages (summarised in Figure 4.8):

1. ***Weights normalisation***: We apply *weights normalisation* to each voxel's time course to down-weight noisy, deep-brain voxels.

2. **Enveloping**: We estimate the oscillatory amplitude envelope at each voxel by taking the absolute of the analytic signal, calculated via the Hilbert transform [92].

3. **Windowed-average temporal down-sampling**: We then temporally down-sample the envelopes at each voxel using a windowed-average. We divide each envelope into non-overlapping[4] windows of length Δ and estimate the mean within each window.

4. **Spatial smoothing**: We spatially smooth the temporally down-sampled envelopes with a 4mm full-width-at-half-maximum Gaussian kernel to improve inter-subject correspondence.

5. **Spatial down-sampling**: We then spatially down-sample the smoothed envelopes to an 8mm grid.

6. **De-meaning**: We remove the mean of each voxel's envelope. This is because the means of the *weights-normalised* envelopes across subjects are distorted by the *weights normalisation* and so do not encode meaningful information (refer to section 5.3).

7. **Concatenation**: We temporally concatenate the preprocessed envelopes from multiple sessions/subjects into a single group matrix.

8. **Dimensionality reduction and pre-whitening**: We use a PCA to select the I largest principle components and normalise each one to have unit variance.

9. **ICA decomposition**: We decompose the pre-whitened data into a set of temporally independent component time courses and associated spatial maps.

10. **Estimation of correlation/covariance maps**: For each time course, we estimate the correlation and covariance between the time course and the concatenated envelopes at each voxel. This gives a correlation map and a covariance map for each component. We normalise each component's covariance map to have unit variance [65].

[4]Using non-overlapping windows has the advantage of greatly reducing the number of samples going into the ICA which helps control the computational cost.

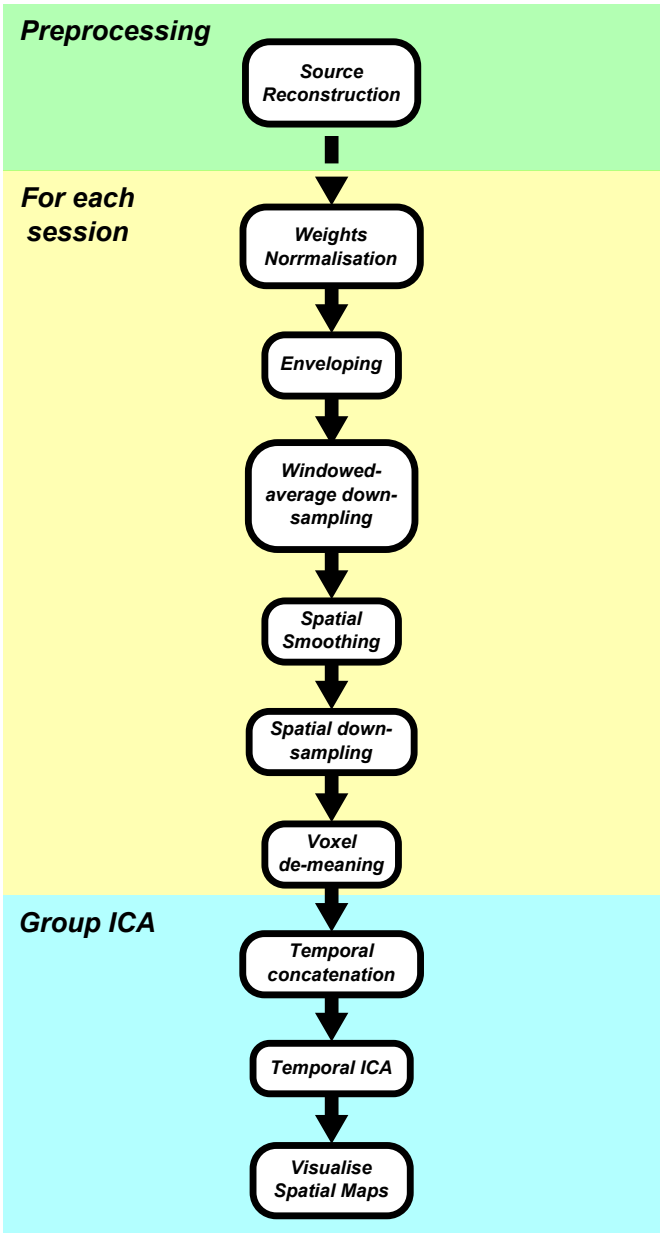


Figure 4.8: *A schematic of the stages involved in performing a group ICA analysis on multi-session source reconstructed data.*

4.5.4 Three studies using group ICA on MEG

In this section, we show the results of three group ICA studies: the first was a validation of the findings of Brookes et al. (2011) by performing an equivalent analysis on an independent resting-state data set acquired on a *CTF* 275-channel system [30]. This allowed us to ascertain if we could replicate their findings and confirm that our analysis framework was performing correctly. The second study was an equivalent group ICA to image RSNs in MEG data acquired on an *Elekta Neuromag*. In particular, we focused on demonstrating that signal space separation (implemented

with *MaxFilter*^{TM}) did not prevent the imaging of RSNs via group ICA. The final study was a simple demonstration to show the importance of *weights normalisation* and de-meaning prior to concatenation.

4.5.4.1 Study 1: Imaging RSNs in *CTF* data

The first application of our ICA framework was to replicate the findings of Brookes et al. (2011) in which they performed two group temporal ICAs on the 1s windowed-average down-sampled envelopes in the α-band (8-13Hz) and β-band (13-30Hz) to extract 8 RSNs (this result is shown in Figure 1.4). Their cohort consisted of 10 subjects, each scanned for 5 minutes with their eyes open and fixating. The data were acquired on a *CTF* 275-channel whole-head system.

We attempted to reproduce this finding on an alternative data set: we used 9 subjects each who has been scanned at eyes-open rest for 10 minutes (refer to section 5.5.1.1 for a full description of the data). We utilised the first 6 minutes of each session in order to have a comparable amount of data. We performed our standard preprocessing pipeline (described in section 3.6), including 3rd order gradiometer correction, ICA de-noising to remove the cardiac, mains and heartbeat artefacts, and visual inspection to identify bad channels and epochs. We beamformed the data (at 6mm resolution) in the α-band (8-13Hz) and β-band (13-30Hz). We note two key methodological differences in our approach compared with Brookes et al.'s analysis. Firstly, we performed sensor-level ICA de-noising whilst they did not. Secondly, our beamformer estimates each optimum dipole orientation as that which projects the maximum power, while Brookes et al. used a SAM beamformer that that estimates the optimum dipole orientation as that which maximises the signal-to-noise ratio (refer to section 2.5.3.4).

Separately for the α-band (8-13Hz) and β-band (13-30Hz) data, we implemented our ICA framework described in section 4.5.3. We used a 1s window length and estimated 25 independent components for each band. We visualised each component by estimating the normalised covariance map between each IC and the input data.

Figure 4.9 shows the findings of our two ICAs. The first column of Figure 4.9

shows Z-statistical images of 8 RSNs derived by applying group spatial ICA to fMRI data [122]. These eight RSNs were used as the gold standard in Brookes et al.'s study [30]. The second column of Figure 4.9 shows equivalent RSNs from our two MEG ICAs. We have found equivalent MEG-derived RSNs for A. the default mode network; B. the frontoparietal networks; C. the sensorimotor network; D. the superior parietal lobule; E. the visual network; F. the anterior cingulate cortex; G. the cerebellum. The default mode network was found in the alpha band, replicating the findings of Brookes et al. We also found the anterior cingulate cortex component in the α-band. All remaining components were found in the β-band (again matching the previous study). One notable difference in our analysis is that we found a single bilateral frontoparietal network.

Overall, this study succeeded in its two aims. We successfully reproduced the findings of Brookes et al. using an independent implementation of the ICA framework for MEG FC analysis and using a separate data set. We do note that our set of RSNs qualitatively looks less like the fMRI RSNs than those found by Brookes et al. (shown in Figure 1.4). We believe this difference could be attributed to one of three causes. 1.) Differences in the data - the results we present do not use exactly the same MEG data. However, we generally find that our analysis produces RSNs with slightly less spatial specificity than those reported by Brookes et al. 2.) ICA de-noising - this is one of the deviations in our analysis. However, we have empirically found that the inclusion of ICA de-noising does not change the spatial maps of our RSNs. 3.) Dipole orientation - it is possible that the SAM beamformer estimate of dipole orientation results in improved sensitivity for RSN detection than our method. Finally, it is possible that a selection of minor differences in the respective MEG analysis pipelines cumulatively lead to Brookes et al.'s implementation offering slightly superior sensitivity.

That said, our result reinforces the finding that RSNs can be robustly extracted from MEG data without any fMRI prior. Furthermore, we have validated our analysis implementation as a tool for imaging RSNs. Combined with the seed-based correlation analysis (Figure 4.5), we have been able to reproduce the spatial and frequency characteristics of resting-state FC that are well-documented in MEG.

4.5.4.2 Study 2: Imaging RSNs in *Elekta Neuromag* data

The analysis described here was presented at the 18th Annual Meeting of the Organization for Human Brain Mapping [87]. It was performed prior to the finalisation of the best-practice preprocessing procedure in chapter 3: the specific analysis used is described in full detail below.

Data and methods: 12 subjects were scanned at rest (eyes-open with a fixation cross) for 6 minutes using an *Elekta Neuromag*. This cohort was a pilot sub-group of the *APOE* cohort described in section 5.5.1.2. Each subject's data was preprocessed without and with signal space separation (implemented via $MaxFilter^{TM}$ [50, 127]). No ICA de-noising was performed. Instead, channels with abnormally high variance were discarded. Furthermore, the data were divided into 15s epochs. Epochs were visually inspected and discarded if found to contain noisy/high variance data. The surviving epochs were concatenated to form a pseudo-continuous data set. Sensors were normalised by their overall variance (rather than their minimum eigenvalue). An LCMV beamformer was used to estimate the source-space activity in three frequency bands (α: 8-13Hz; β: 13-30Hz; γ: 30-70Hz).

The envelopes of *weights-normalised* activity for each band were estimated and down-sampled using a windowed-average down-sampling (1s window length). The envelopes were spatially smoothed, spatially down-sampled and de-meaned before being temporally concatenated. Temporal group ICA was performed separately on each frequency band to give 25 independent components. For each component, the map of correlation scores between the IC time course and the concatenated envelopes at each voxel were estimated.

Results and discussion: Figure 4.10 shows the results of the three ICAs on the data with and without SSS preprocessing. 11 RSNs have been identified in both the SSS'd and non-SSS'd data. In the left column we show the fMRI Z-statistical maps of RSNs taken from [122]. In the middle column, we show the equivalent RSN correlation maps found in the ICAs performed on the non-SSS'd data. In the right

column, we show the equivalent RSN correlation maps shown in the SSS'd data. In the α-band we found a component corresponding to the default mode network. In the β-band, we found components corresponding to the right frontoparietal network, left frontoparietal network, sensorimotor network, somatosensory network, superior parietal lobule, anterior cingulate cortex, cerebellum, and visual network. In the γ-band, we found components corresponding to the lateral visual network and the medial visual network.

This analysis has three important outcomes. Firstly, it is an additional validation of the findings by Brookes et al. in 2011 [30]. Our previous analysis in section 4.5.4.1 utilised data from the same *CTF* scanner as [30] and also contained subjects who were scanned as part of the original study. This analysis is completely independent: the type of scanner and the cohort is different. Secondly, this is the first demonstration of imaging RSNs using an *Elekta Neuromag* system.[5] Thirdly, we have demonstrated that signal space separation can be applied as part of the preprocessing of *Elekta Neuromag* resting-state data. SSS (via *MaxFilter*^{TM}) is neither essential nor does it prevent the imaging of RSNs so long as the reduced dimensionality of the *MaxFiltered* data is taken into account (refer to section 3.5.1 for a full discussion of this issue).

4.5.4.3 Study 3: Demonstrating the importance of *weights normalisation*

Data and methods: We acquired 10 minutes of eyes-closed resting-state and 10 minutes of visual active-state (where the participants watched a silent movie) for 8 subjects. Data were acquired using a *CTF* 275-channel whole-head system. For a full description of the data set, refer to section 5.5.1.1. We performed 3rd order gradiometer correction to remove far-field interference, sensor-space ICA de-noising to remove cardiac, eye-blink and mains artefacts, and manual identification of noisy channels and epochs. We band-pass filtered the data into the β-band (13-30Hz) and estimated the source-space activity across a 6mm grid using an LCMV beamformer. Refer to section 3.6 for a full description of the preprocessing stages. In source-space, we esti-

[5] At the time of writing, we performed a search of the SCOPUS database for the following terms: (neuromag OR elekta) AND (resting state OR rsn). We found one result looking at sensor-level differences in resting-state oscillatory power in HIV positive populations [12].

mated the oscillatory amplitude envelope and performed a windowed-average down-sampling using 2s non-overlapping windows. We performed three different group concatenations:

- ***Weights-normalised* and de-meaned**: We applied *weights normalisation* to every voxel's time course and removed the mean of the *weights-normalised* envelope before temporally concatenating all the sessions.
- **No *weights normalisation***: We did not apply the *weights normalisation* correction prior to envelope estimation.
- **No de-meaning**: We did not subtract the mean of the *weights normalised* envelopes prior to concatenation.

For each concatenated envelope matrix, we performed a temporal ICA into 25 temporally independent components. For each component, we estimated the normalised covariance map.

Results and discussion: Figure 4.11 shows three covariance maps for each of the ICAs (A. *Weights-normalised* and de-meaned, B. No *weights normalisation* and C. No de-meaning). For each ICA, we selected the components that best matched the sensorimotor network (SMN) and the left and right frontoparietal networks (FPN). Only when both *weights normalisation* and de-meaning were applied did we find components showing bilateral functional connectivity in the SMN and anterior-posterior connectivity in both FPNs. As Figure 4.11 shows, not applying *weights normalisation* or de-meaning does not completely destroy the resulting RSNs but does noticeably degrade them.

4.6 Summary

In this chapter we presented three sequential bodies of work that build upon each other to produce a data driven framework for robustly extracting RSNs from resting-state MEG data. The first section presented an investigation into the optimum *carrier* and

co-variation frequencies using pairs of nodes from known functionally-connected networks. We found that the optimum down-sampling window length was between 1-4s (corresponding to a low-pass filter cut-off of approximately 0.5Hz). This analysis gave an initial insight into the time scales of the neuronal processes that drive functional connectivity and more importantly informed the choice of *co-variation frequency* for all subsequent FC analyses.

The second section was an exposition of the seed-based correlation analyses used in this thesis. Although ICA has been adopted as the primary technique for identifying functionally-connected networks, seed-based correlation has been an invaluable tool for validating most stages of our analysis.

Finally, we presented a framework for using group ICA to discover resting-state networks, including a demonstration of some of the pitfalls in the analysis. In this chapter and chapter 3 we have invested a lot of effort in developing and validating an analysis pipeline that can be applied to both *CTF* and *Elekta Neuromag* data. This effort has culminated in two reproductions of the work by Brookes et al. showing the simultaneous extraction of multiple RSNs from MEG data [30].

Chapters 5 and 6 build on the ICA framework developed in this chapter by outlining statistical frameworks for inferring changes in functional connectivity across populations and during a task respectively.

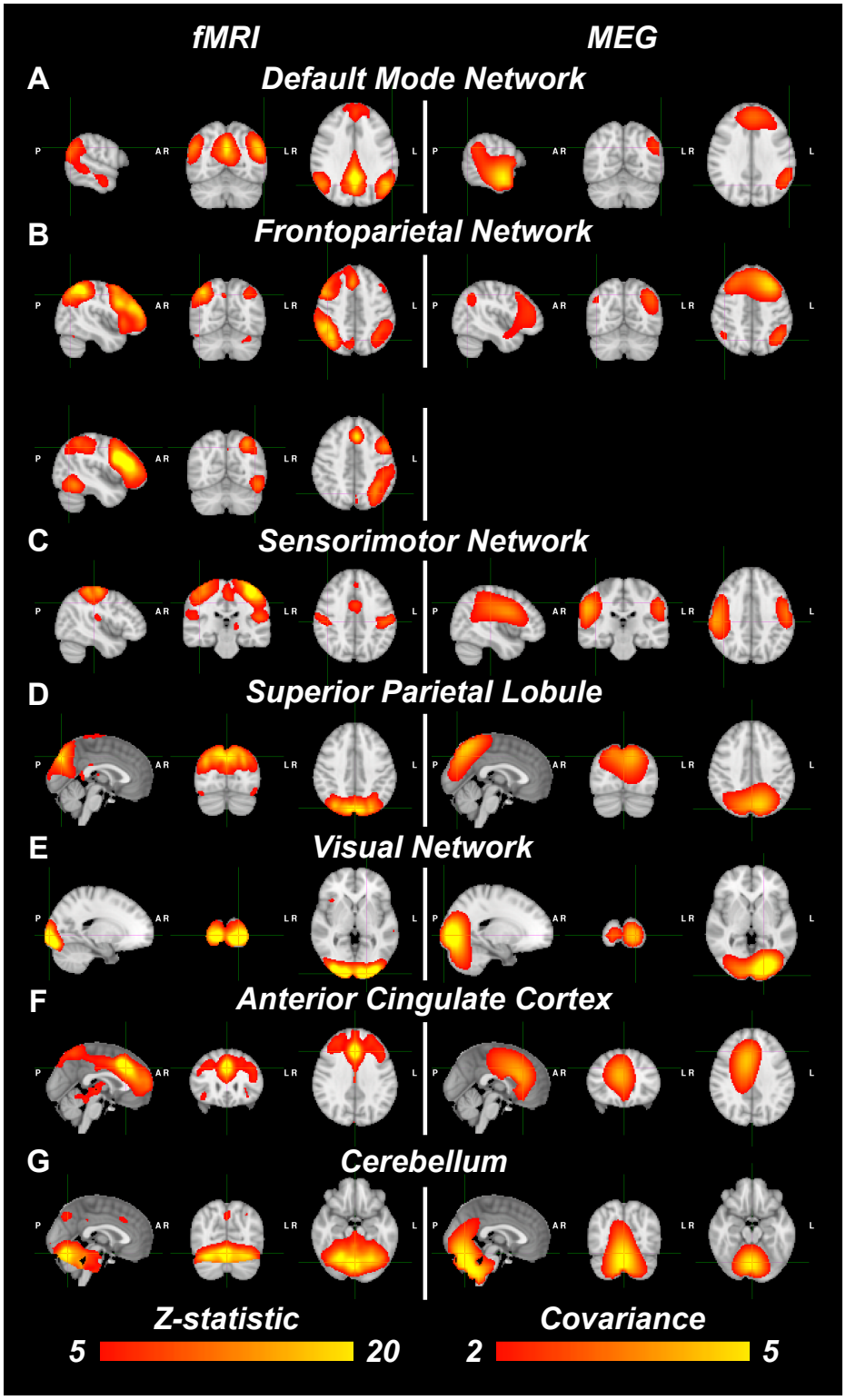


Figure 4.9: *Replication of the MEG-derived RSNs from [30]. We performed two ICAs: on the α-band and β-band envelopes (down-sampled with 1s non-overlapping windows. For each ICA, we estimated 25 temporally independent components and plotted their normalised covariance maps (right). Alongside we have displayed the equivalent fMRI maps (left) taken from [122]. We found 7 RSNs: A. the default mode network; B. the frontoparietal networks; C. the sensorimotor network; D. the superior parietal lobule; E. the visual network; F. the anterior cingulate network; G. the cerebellum. A and F were taken from the α-band ICA. The remaining networks came from the β-band ICA.*

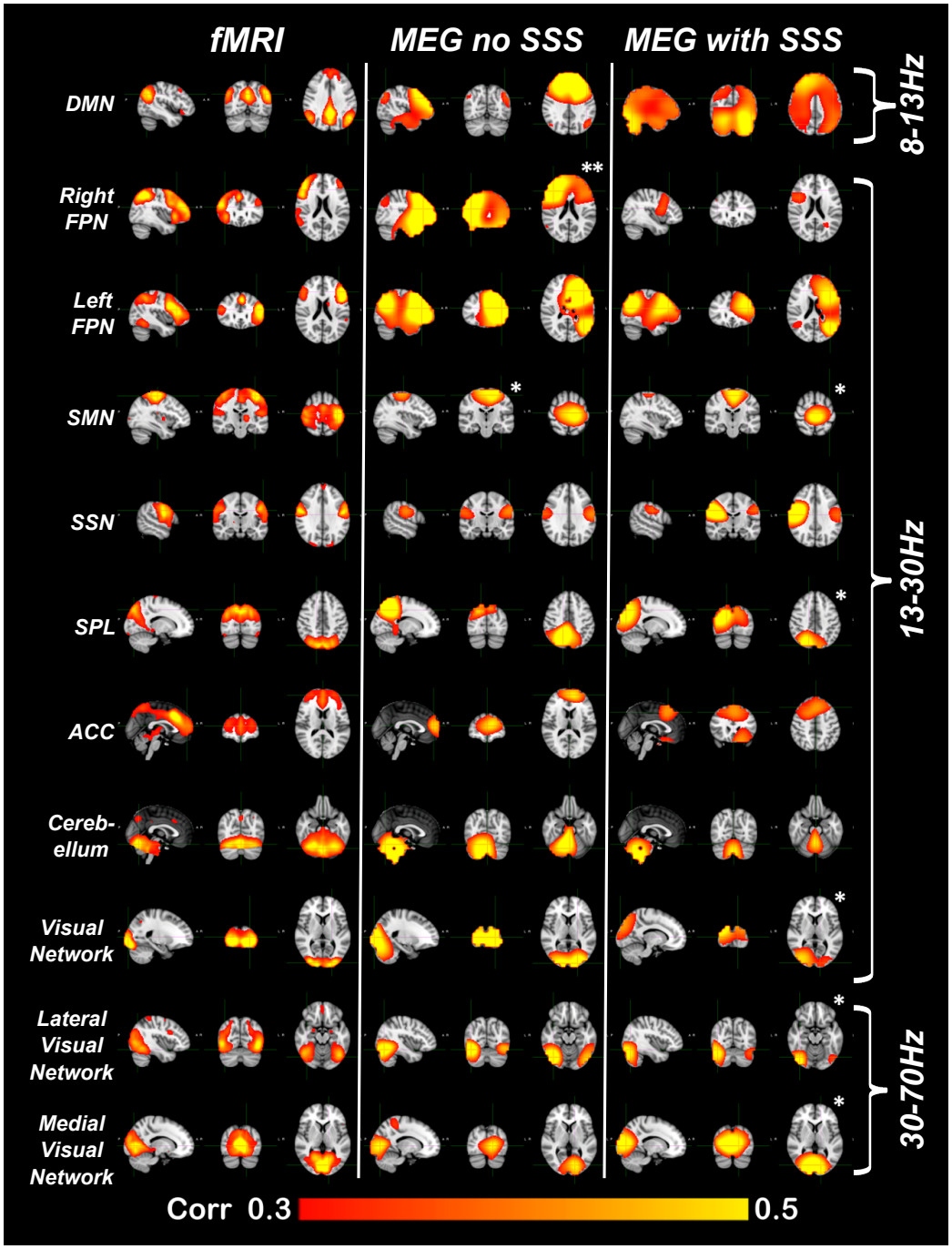


Figure 4.10: *The results of the three ICAs on the data with and without SSS preprocessing. 11 RSNs have been identified in both the SSS'd and non-SSS'd data. In the left column we show the fMRI Z-statistical maps (thresholded at $Z = 3$) of RSNs taken from [122]. In the middle column, we show the equivalent RSN correlation maps found in the ICAs performed on the non-SSS'd data. In the right column, we show the equivalent RSN correlation maps extracted from the SSS'd data. In the α-band we found a component corresponding to the default mode network (DMN). In the β-band, we found components corresponding to the right frontoparietal network (FPN), left frontoparietal network, sensorimotor network (SMN), somatosensory network (SSN), superior parietal lobule (SPL), anterior cingulate cortex (ACC), cerebellum; visual network. In the γ-band, we found components corresponding to the lateral visual network and medial visual network. All correlation values are between at 0.3 and 0.5 except: *correlation values are between 0.5 and 0.7; **correlation values are between 0.25 and 0.5.*

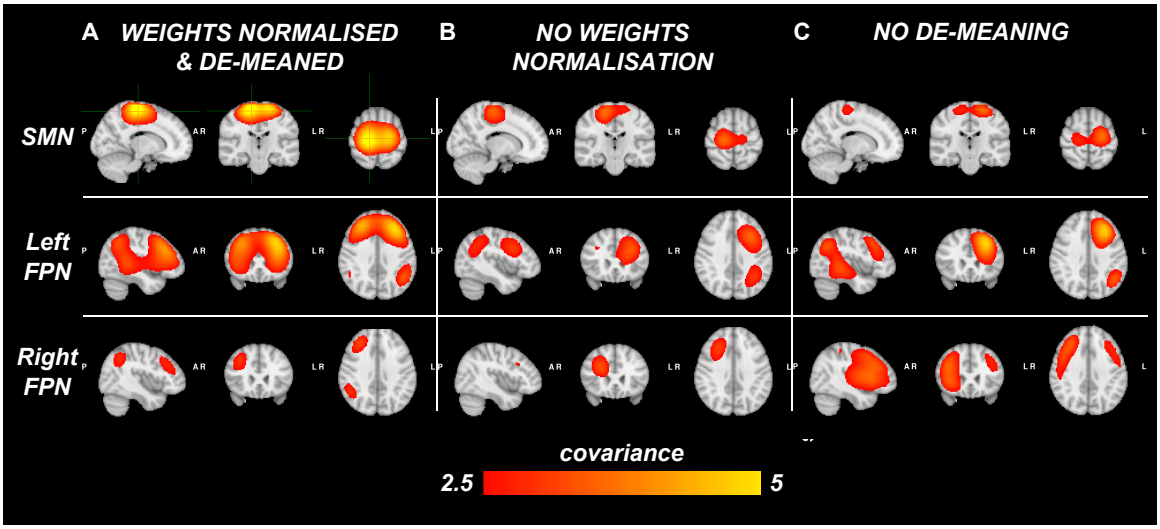


Figure 4.11: *A demonstration of the importance of beamformer weights normalisation and de-meaning of the weights-normalised envelopes prior to temporal concatenation for group ICA. Three ICAs were performed on the concatenation of 2s windowed-averaged envelopes of β-band data from 8 sessions of eyes-closed rest and 8 sessions of visual active-state. For each ICA, we selected the covariance maps of the components corresponding to the sensorimotor network and the left and right frontoparietal networks. A. Temporal concatenation performed with weights normalisation and de-meaning. B. Without weights normalisation. C. Without de-meaning.*

Chapter 5

Investigating group differences in resting-state networks with MEG

5.1 Chapter abstract

In this chapter, we show, both theoretically and empirically, that beamformer *weights normalisation* is a major confound for multi-session statistics in MEG and must be accounted for accordingly. We go on to present a novel framework, based on dual regression, for evaluating differences in functional networks across different populations. We demonstrate that MEG-adapted dual regression (MADR) can utilise **any** group average spatial basis set and determine sub-group differences relative to that average. We use MADR to show that carriers of the APOE-ϵ4 gene have significantly lower variance of oscillatory power associated with left and right frontoparietal networks, when compared with normal APOE-ϵ3 homozygotes.

5.2 Introduction

Analysis of resting-state brain activity has led to the discovery of a range of spatially distributed but functionally integrated resting-state networks (RSNs) which can be imaged using multiple modalities [13,30,112]. In chapter 4, we demonstrated a framework for applying temporal ICA to extract a set of RSN spatial maps from MEG data

(without any prior information from other modalities).

One interpretation of these RSN maps is as an efficient, low-dimensional representation of the high-dimensional voxel-level data. The RSNs can be treated like a functionally-derived spatial basis set onto which our resting-state data can be projected. It has been shown that the RSN spatial basis set can be linked to regions associated with a range of cognitive functions and that differences in RSNs can be detected and used as biomarkers for differences in functional connectivity across different populations (such as carriers of the APOE-ϵ4 risk gene for AD) [55, 122, 129]. This is why RSNs are such a widely investigated spatial basis set representation of brain function.

An important challenge is how to appropriately test for inter-session/subject differences in resting-state function, for example, between different populations, when provided with **any** general spatial basis set. The spatial basis set could correspond to ICA spatial maps derived from the neuroimaging data in question (where the spatial maps correspond to the group average of that data set), or from independent sources such as a standard atlas, or an independent neuroimaging dataset. Whatever the spatial basis set, we would like a technique that can take a single group average set of maps as a starting point and infer the maps specific to the sub-groups that make up our cohort in question. Dual regression is a framework that has been previously developed for this purpose, originally for estimating subject-specific maps from the output of spatial ICA performed on concatenated multi-subject fMRI data [14]. It uses two stages of multiple regression to estimate the subject-specific time courses, followed by the subject-specific spatial maps. Standard mass univariate group-level statistical techniques, such as the General Linear Model, can then be used to infer sub-group differences.

In this chapter, we show that specific care must be taken with regards to the beamformer *weights normalisation* to ensure that inter-session/subject effects are comparable when each session/subject is beamformed separately. We demonstrate the effect of the *weights normalisation* confound through a theoretical comparison of the true variance of a neuronal dipole and the beamformer estimate of the variance.

We also demonstrate, using a simulation and using real MEG data, that ignoring the *weights normalisation* confound can have serious repercussions (such as reversing the direction of the effect being measured) for multi-session statistical analyses, both in simple voxel-wise analyses and in dual regression analyses.

We then proceed to adapt dual regression for MEG to produce a framework for performing group-level statistics on the low frequency envelope oscillations of the band-limited time courses of beamformed MEG data using general spatial basis sets derived from approaches such as group-ICA. In order to determine the sensitivity of dual regression to different types of group difference, we use a suite of simulations, based on the ICA mixing model, to generate specific group differences and then use our MEG-adapted dual regression (MADR) analysis to recover those differences. We validate the MADR analysis by using it to detect increased oscillatory alpha power in visual RSNs during eyes-closed rest compared with continuous visual stimulation. Finally, we demonstrate the use of the MADR framework to investigate differences in resting-state oscillatory brain activity in carriers of the APOE-ϵ4 gene, which has been proven to increase an individual's risk of developing Alzheimer's disease.

5.3 Comparing variances across multiple sessions of beamformed data

Here, we build on the theoretical formulation of the beamformer that we presented in section 2.5.3. We focus on how to compare variances across separately beamformed sessions. We introduce the *weights normalisation* confound which arises as a consequence of the fact that *weights-normalised* beamformed data gives a biased estimate of the true variance of a dipole. The concepts presented here are essential for MADR where our aim is to compare networks of co-varying band-limited power across a cohort of subjects for which each session of data is source-reconstructed separately. However, it should be noted that the *weights normalisation* confound will effect **any** statistical analysis that compares variances (or other measures that are dependant on variance) across multiple sessions of data that have been beamformed separately.

5.3.1 Deriving the *weights normalisation* confound

In section 2.5.3.3 we derived expressions for the raw (*non-weights-normalised*) beamformer estimate of the variance, $\hat{\sigma}^2_{q'}$, at a voxel, q', using weights $\mathbf{W}_{q'}$, in terms of the true variance, ${\sigma_{q'}}^2$, plus contributions due to signal leakage and the projection of sensor noise:

$$\hat{\sigma}^2_{q'} = \underbrace{{\sigma_{q'}}^2}_{\text{true variance}} + \underbrace{\mathbf{W}_{q'}(\sum_{q \neq q'} \mathbf{H}_q {\sigma_q}^2 \mathbf{H}_q^T)\mathbf{W}_{q'}^T}_{\text{signal leakage}} + \underbrace{{\sigma_{\boldsymbol{\varepsilon}}}^2 \mathbf{W}_{q'}\mathbf{W}_{q'}^T}_{\text{noise projection}} \tag{5.1}$$

and the *weights-normalised* estimate of the variance, $\tilde{\sigma}^2_{q'}$:

$$\tilde{\sigma}^2_{q'} = \frac{{\sigma_{q'}}^2}{\mathbf{W}_{q'}\mathbf{W}_{q'}^T} + \frac{\mathbf{W}_{q'}(\sum_{q \neq q'} \mathbf{H}_q \sigma_q^2 \mathbf{H}_q^T)\mathbf{W}_{q'}^T}{\mathbf{W}_{q'}\mathbf{W}_{q'}^T} + {\sigma_{\boldsymbol{\varepsilon}}}^2 \tag{5.2}$$

If we consider any analysis that attempts to compare variances across multiple beamformer sessions we can demonstrate that we must use the *non-weights-normalised* estimates of the variance, $\hat{\sigma}^2_{q'}$. Consider two sessions of separately beamformed data, *1* and *2*, between which we intend to test for differences in variance (e.g. to see if we can detect a change in oscillatory power in a particular frequency band). At the q'^{th} voxel, we have *non-weights-normalised* estimates $\hat{\sigma}^2_{q',1}$ and $\hat{\sigma}^2_{q',2}$. The difference between these estimates is:

$$\hat{\sigma}^2_{q',1} - \hat{\sigma}^2_{q',2} = {\sigma_{q',1}}^2 - {\sigma_{q',2}}^2 + e... \tag{5.3}$$

where $e...$ represents the additional error terms (i.e. the contributions to our variance estimate due to signal leakage and projected sensor noise). We assume that these contributions are approximately equal across *sessions 1* and *2*. Most importantly, this assumption requires that the noise variance of the scanner is constant across the two sessions. If we perform our statistics on the *weights-normalised* variances, $\tilde{\sigma}^2_{q',1}$ and $\tilde{\sigma}^2_{q',2}$, equation 5.2 tells us that we get a biased estimate of the difference in

variance:

$$\tilde{\sigma}^2_{q',1} - \tilde{\sigma}^2_{q',2} = \frac{{\sigma_{q',1}}^2}{\mathbf{W}_{q',1}\mathbf{W}^T_{q',1}} - \frac{{\sigma_{q',2}}^2}{\mathbf{W}_{q',2}\mathbf{W}^T_{q',2}} + \tilde{e}... \tag{5.4}$$

where $\tilde{e}...$ represents the additional error terms (due to signal leakage and the projection of sensor noise) with their respective session specific *weights normalisation* applied. This concept has been previously established in Vrba et al.'s pseudo-T-statistics for comparing active versus control blocks within a single session of data [136]. However, we have restated it as it is critical for correctly interpreting statistics performed on the envelopes from different sessions that have been separately beamformed. Using a simulation we can show that, under certain conditions, the *weights normalisation* confound can actually reverse the measured direction of an effect of interest.

Consider a single dipole A across two separate sessions, *1* and *2* where the leadfields do not change between sessions. Let us assume the dipole actually has higher standard deviation in *session 1* than in *session 2* (i.e. the ground truth is $\sigma_{A,1} = \alpha \times \sigma_{A,2}$ where $\alpha > 1$). Now consider our *weights-normalised* beamformer estimates of the standard deviation of *dipole A*, $\tilde{\sigma}_{A,1}$ and $\tilde{\sigma}_{A,2}$. We can show using equation 5.2 that:

$$\tilde{\sigma}_{A,1} = \frac{\sigma_{A,1}}{\sqrt{\mathbf{W}_{A,1}\mathbf{W}^T_{A,1}}} + \tilde{e}_1... \tag{5.5}$$

where $\mathbf{W_{A,1}}$ is the beamformer weights vector for *dipole A* in *session 1* and $\tilde{e}_1...$ is short-hand for the *weights-normalised* error terms corresponding to the signal leakage and noise projection contributions. We can find similar expressions for *session 2*. Ignoring the error terms $\tilde{e}_1...$ and $\tilde{e}_2...$ for the moment, we can see that when

$$\sigma_{A,1} = \alpha \times \sigma_{A,2} \tag{5.6}$$

then

$$\tilde{\sigma}_{A,1} < \tilde{\sigma}_{A,2} \tag{5.7}$$

if and only if

$$\sqrt{\mathbf{W_{A,1}}\mathbf{W_{A,1}}^T} > \alpha\sqrt{\mathbf{W_{A,2}}\mathbf{W_{A,2}}^T} \tag{5.8}$$

where α is the proportional change in variance and $\alpha > 1$. Looking at the *weights normalisation* factor, $\sqrt{\mathbf{W_{A,1}}\mathbf{W_{A,1}}^T}$, of the weights vector, $\mathbf{W_{A,1}}$ for *dipole A* in *session 1*, it is mathematically equivalent to the Euclidean length. The simplest situation in which we can observe a change in the beamformer weights vector for *dipole A* between *sessions 1* and *2* is when we have at least two additional sources whose variances change by different proportions between the two sessions. In this case, it is possible to generate two different sets of beamformer weights vectors for *dipole A*, whose proportional change in length is enough to swamp the genuine change in the variance of *dipole A*.

5.3.2 Simulating the *weights normalisation* confound

We demonstrate this with a very simple 2-dimensional simulation involving three dipoles (*A*, *B* and *C*) in a 2-sensor MEG system. Figure 5.1A is a schematic showing the approximate geometry of this simulation. Figure 5.1B shows the lead-field vectors (constant across *sessions 1* and *2*) for the three dipoles. For simplicity, we have made all the lead-field vectors unit magnitude (although this does not affect the generality of our argument). We generated three orthogonal, normally distributed time courses for the three dipoles for each session. Figure 5.1C shows the standard deviations of the dipoles. In our simulation, *dipole A* shows a small reduction in standard deviation between *sessions 1* and *2*, *dipole B* shows a large reduction, and *dipole C* remains unchanged. We projected both sessions of data through our simulated lead-fields and then separately beamformed each session. Figure 5.1D shows the difference in standard deviation between *session 1* and *2* for *dipole A* for the ground truth, the *non-weights-normalised* estimate, and the *weights-normalised* estimate. Note that

under these conditions, our *weights-normalised* estimate of the standard deviation in *session 2* is bigger not smaller than in *session 1*.

We can see why by considering the weights vectors, $\mathbf{W_{A,1}}$ and $\mathbf{W_{A,2}}$, from *sessions 1* and *2* for *dipole A*, plotted as dashed cyan and magenta vectors alongside the lead-field vectors in Figure 5.1B. The unity pass band constraint limits the weights vectors to a specific subspace. In *session 1*, *dipole B* has the greatest variance. Consequently, the beamformer tries to minimise the projected power and so finds the optimum weights to be almost orthogonal to the lead-fields of *dipole B*. In *session 2*, *dipole B*'s variance greatly reduces and the beamformer adapts by finding a set of weights that are now more orthogonal to the lead-fields of *dipole C*, whose projected variance is relatively greater in *session 2*. This has the unintended consequence of changing the Euclidean length of the weights vectors by a greater proportion than the change in standard deviation of *dipole A*. As such the *weights-normalised* estimate of the change in standard deviation for *dipole A* actually reverses sign.

5.3.3 The *weights normalisation* confound in real MEG data

This change in the valence of the estimated difference in standard deviation or variance between sessions is not limited to simulations but can be observed in real MEG data. Here, we show such an example. We acquired 10 minutes of *eyes-closed* rest and 10 minutes of visual *active-state* data for 8 subjects. For every session, we estimated the oscillatory envelope of the alpha band (8-13Hz) beamformed neural activity over a 6mm grid. Full details of the *eyes-closed*/*active-state* data set and analysis are given below in sections 5.5.1.1 and 5.5.2 respectively.

We performed a voxel-wise analysis in order to demonstrate the bias in variance estimation that *weights normalisation* introduces. Eyes-closed resting-state has an extremely consistent increase in alpha power compared with eyes-open states [20,33]. Specifically, we applied a paired t-test between *eyes-closed* and *activate-state* sessions (via a GLM) to the mean and variance of the envelope at each voxel over all subjects. Performing statistics on the mean of the envelope for a given voxel is equivalent to testing for changes in the mean of oscillatory power for that voxel/frequency band.

Similarly, performing statistics on the variance of the envelope for a given voxel is equivalent to testing for changes in the variability of oscillatory power for that voxel/frequency band. We did this test using both the *non-weights-normalised* and *weights-normalised* envelopes to demonstrate the confound that *weights normalisation* introduces.

These maps are shown in Figure 5.2. Without *weights normalisation*, we found a significant ($p_{corrected} \leq 0.05$) increase in the mean and variance of the alpha band envelope in the visual cortex. However, when *weights normalisation* was applied, we observed the reverse: a significant decrease in alpha power in *eyes-closed* sessions relative to the *active-state* sessions. In addition, the spatial location of this power decrease was deeper and more medial. Multiple comparisons were accounted for using threshold-free cluster enhanced (TFCE) permutation testing in **FSL** (*FMRIB, Oxford*) and all super-threshold voxels in Figure 5.2 survived correction.

5.3.4 Discussion

The essential take-home point of this section is that beamformer *weights normalisation* can completely distort the absolute estimate of variance when performing multi-session statistics. We demonstrated this point in three ways: theoretically, with a simulation, and with a practical demonstration of the effect in real data.

An important question is whether *weights normalisation* by definition reverses the effect direction or whether our voxel-wise analysis is an extreme, (potentially) worst case scenario. Our simple 2D simulation shows that the reversal of inferred effect direction is not an automatic consequence of *weights normalisation* but depends on the specific relative changes in variance of the dipoles and their lead-field orientations.

The mechanism by which we simulated a change in effect valence could plausibly be occurring in our *eyes-closed*/*active-state* analysis. If we consider the medial visual cortex where we observed a reversal in the direction of alpha power change, it is possible that this region experienced a small increase in alpha power during the *eyes-closed* rest compared with the *active-state*. However, other visual areas may have experienced much larger alpha power increases in the *eyes-closed* condition compared

with the *active-state*. Furthermore, it is likely that other brain areas showed no difference in alpha power between the two conditions. As such, this experiment could quite plausibly be operating in a similar regime as our simulation (be it in 275 dimensions rather than 2).

We developed this argument as it is essential for correctly inferring group statistics on the oscillatory envelopes of neural activity that are used in functional connectivity analyses. We made a key assumption in our analysis of the *weights normalisation* confound. This was that the contributions to the beamformed estimate of variance due to signal leakage and sensor noise were approximately constant. It is reasonable to assume that sensor noise is approximately constant (particularly in our analysis where the *eyes-closed* and *active-state* scans were acquired in a single scanning session). However, the signal leakage contribution is a function of the relative variances of all the neuronal sources. It is expected that these variances will be different between sessions and subjects. As such, separate work is needed to understand the effect of inter-session variability in signal leakage to variance-based analyses over multiple sessions. Finally, this issue could extend beyond resting-state analysis and have implications for the correct way to perform group-level statistics on beamformed, task-positive, multi-session data.

5.4 Adapting dual regression for MEG data

In the previous section, we demonstrated that it is essential **not** to compare *weights-normalised* estimates of variance across multiple separately beamformed sessions. Using this insight, we now present an adaptation of dual regression which allows us to compare multi-session, beamformed data in a principled manner.

Given a group-level spatial basis set, $\tilde{\mathbf{A}}$ (which is Q voxels by I components), we would like to infer if there are any sub-group differences relative to that basis set.[1] To achieve this, we must estimate the session-specific spatial basis sets, which can be done using dual regression [14]. Commonly, in fMRI dual regression, the spatial

[1] We use a tilde to signify that this basis set does not model the depth bias of *non-weights-normalised* MEG data.

basis set is estimated by performing spatial ICA on the temporally concatenated fMRI data. Given such a group level basis set, dual regression estimates the group-level concatenated component time courses, $\mathbf{S}_{concat}$ (which is I components by $T \times N$ samples, where T is the average number of samples per subject and N is the number of subjects), via a spatial regression:

$$\mathbf{S}_{concat} = \tilde{\mathbf{A}}^{+}\mathbf{B}_{concat} \tag{5.9}$$

where $\tilde{\mathbf{A}}^{+}$ is the [I components by Q voxels] pseudo-inverse of the spatial basis set matrix and $\mathbf{B}_{concat}$ is the [Q voxels by $T \times N$ samples] matrix of zero-mean, BOLD time series at every voxel, temporally concatenated over subjects. Using the segment of the group-level component time courses specific to the n^{th} session, we can estimate the [Q voxels by I components] spatial maps, $\tilde{\mathbf{A}}_n$, for that session via the following temporal regression.

$$\tilde{\mathbf{A}}_n = \mathbf{B}_n{\mathbf{S}_n}^{+} \tag{5.10}$$

where ${\mathbf{S}_n}^{+}$ is the [T samples by I components] pseudo-inverse of the component time courses specific to the n^{th} session and $\mathbf{B}_n$ is the [Q voxels by T samples] segment of BOLD data from the n^{th} session.

However, in MEG we have the following issue: in order to perform correct univariate statistics at each voxel, we must use the *non-weights-normalised* envelopes in order to have an unbiased estimate of the contribution from the true variance (as we showed in section 5.3). However, these envelopes contain the depth related spatial bias (which can be seen in Figure 2.4). Most spatial basis sets will not model this spatial bias and so the initial spatial regression will yield poor estimates of the group concatenated time courses. Our solution to this is to perform the spatial regression using the *weights-normalised* envelopes $\tilde{\mathbf{E}}$ (so that the envelope data does not have the depth bias and matches the spatial basis set), and then the temporal regression

using the *non-weights-normalised* envelopes $\hat{\mathbf{E}}$ (so that we have an unbiased estimate of the variances). The resulting session-level spatial maps will have an unbiased estimate of the true variance. Henceforth, we will refer to this variant of dual regression as *MEG-adapted dual regression* (MADR). Explicitly, we use the following steps:

MADR stage 1:

$$\mathbf{S}_{concat} = \tilde{\mathbf{A}}^{+}\tilde{\mathbf{E}}_{concat} \tag{5.11}$$

where the columns of $\tilde{\mathbf{A}}$ have been forced to zero mean and unit variance prior to estimating the pseudo-inverse.

MADR stage 2:

For the n^{th} session:

$$\hat{\mathbf{A}}_n = \hat{\mathbf{E}}_n {\mathbf{S}_n}^{+} \tag{5.12}$$

where the rows of $\hat{\mathbf{E}}_n$ and $\mathbf{S}_n$ have had their means removed (as the means cannot be mapped correctly between the component- and voxel-levels by the multiple regression) and the rows of $\mathbf{S}_n$ have been normalised to unit variance prior to estimating the pseudo-inverse (so that all the variance information is contained in the spatial maps $\hat{\mathbf{A}}_n$). Figure 5.3 graphically shows the two stages of MADR.

We can then perform standard univariate statistics at every voxel of the session-specific spatial maps to find spatially localised group differences, in the manner of standard dual regression. Furthermore, we can perform similar statistical tests on the variance of the component time courses associated with each spatial map. To estimate each component time course's variance, we simply estimate the variance of each column of $\hat{\mathbf{A}}_n$, which corresponds to the variance of each spatial map. This is because we forced the session-specific component time courses all to have unit variance in MADR stage 2, shifting the time course variance into the spatial maps.

By projecting the voxel-level means through the session-specific maps, we can also generate an estimate of the mean of the component time course for each session and

perform similar univariate statistics. To estimate the mean of each component time course, $\boldsymbol{\mu}_{\mathbf{S}_n}$, we project $\boldsymbol{\mu}_{\hat{\mathbf{E}}_n}$, the vector of means of the rows of $\hat{\mathbf{E}}_n$, through the pseudo-inverse of the estimate of the session-specific maps, $\hat{\mathbf{A}}_n^+$:

$$\boldsymbol{\mu}_{\mathbf{S}_n} = \hat{\mathbf{A}}_n^+ \boldsymbol{\mu}_{\hat{\mathbf{E}}_n} \tag{5.13}$$

In summary, MEG-adapted dual regression (MADR) provides three possible tests: on the voxel-level values of the spatial maps, the variances of the component time courses and the means of the component time courses. To clarify what these tests mean, consider an analysis where we use a canonical spatial basis set consisting of RSN maps (taken from a gold standard fMRI analysis). The statistics performed on the component means will find group differences in the average band-limited power associated with each network. The statistics performed on the variance of the component time courses will find group differences in the variance of the band-limited power associated with each network. Finally, the voxel-wise mass univariate statistics will find spatially localised group differences in the variance of the band-limited power associated with each network.

5.4.1 Choice of spatial basis set

We have so far considered using a generic spatial basis set as the starting point for MADR. Here, we consider several possible options for the basis set. In the context of resting-state network analysis, a set of RSN maps estimated via spatial group ICA performed on the concatenated BOLD time series is an obvious starting point. Using fMRI-derived spatial maps has the advantage of being directly linked to the more established consensus definition of the main resting-state networks [122]. Alternatively, an atlas-based spatial basis set or an anatomically defined parcellation might be more appropriate for some analyses. Parcellations correspond to single region-of-interest analyses.

A spatial basis set can be estimated directly from the MEG data itself by performing a temporal ICA decomposition on the group concatenated envelopes [26, 30, 88].

In section 4.5, we developed such a framework for performing temporal ICA on the envelopes of source space MEG data. A MEG-derived basis set has the advantage of being drawn from the data on which the dual regression analysis is performed. As we showed in section 4.5.4, it is essential to use the *weights-normalised* envelopes in the ICA decomposition as *weights normalisation* down-weights the noisy, deep brain voxels. This gives an unbiased estimate of the noise at each voxel, which greatly improves the performance of the ICA [15]. See Figure 4.11 for a demonstration of the importance of *weights normalisation* prior to the ICA decomposition. Accordingly, MADR is necessary even if the basis set is estimated directly from the MEG data.

5.5 Validation of MEG-adapted dual regression

In order to validate our MADR framework, we performed three analyses: 1. Simulations to investigate what types of group difference could be detected with MADR. 2. Application of MADR to detect differences in RSNs in *eyes-closed* rest compared with a visual *active-state*. 3. Application of MADR to detect differences in RSNs between carriers of the APOE-ϵ4 gene and ϵ3 homozygotes.

5.5.1 Data

Three data sets were used in this chapter.

1. **Eyes-closed/active-state cohort**: For development and validation, a very simple but well characterised paradigm was used where each participant was scanned under three conditions: 1.) *eyes-closed* resting-state; 2.) eyes-open resting-state; 3.) eyes-open watching a continuous movie - which we refer to as the visual *active-state*. Although all three states were acquired, in this analysis we only considered the contrast between the *eyes-closed* and *active-state* sessions (selected as it should produce the largest effect). The eyes-open resting-state data was used in the replication of Brookes et al.'s ICA study in section 4.5.4 (Figure 4.9) [25]. We also used the combined *eyes-closed/active-state* data

to demonstrate the importance of *weights normalisation* and de-meaning prior to performing group ICA (section 4.5.4, Figure 4.11).

2. **APOE cohort**: We also applied the MADR framework to a resting-state data set of participants who had been grouped by genotype (specifically by their alleles of the APOE gene) in order to assess if the framework can detect changes in functional connectivity due to genotype, similar to those found by Filippini et al. and Trachtenberg et al. [55, 129].

3. **Simulated cohort**: We also produced a set of simulated data for interrogating what flavours of group difference can be detected by the MADR analysis. This simulated cohort matched the *eyes-closed/active-state* cohort in group sizes (8 simulated subjects, each with 2 sessions), thus approximately matching the statistical power.

5.5.1.1 *Eyes-closed/active-state cohort*

Participants: Ten healthy volunteers were recruited. The cohort comprised 7 males (all right-handed) and 3 females (2 right-handed) with a mean age of 27 years and standard error of 0.48 years. The study was approved by the University of Nottingham Medical School Research Ethics Committee.

Data acquisition: Each participant underwent a 30 minute scan with three blocks, each lasting 10 minutes. Each participant was scanned in a supine position.

- **Session 1** - *eyes-closed* rest: The participant was at rest with their eyes closed and instructed not to fall asleep.

- **Session 2** - eyes-open rest: During this block the subject opened their eyes and fixated on a cross.

- **Session 3** - visual *active-state*: During this block, the participants watched a movie, which was projected through a waveguide in the magnetically shielded room onto a screen placed 40cm in front of the subject.

In this chapter, the eyes-open resting state sessions were discarded. MEG data were acquired using a *CTF* 275 channel whole-head system. The data were acquired at a sampling frequency of 600Hz and synthetic 3rd order gradiometer correction was applied to reduce external interference. Head localisation within the MEG helmet was achieved using three electromagnetic head position indicator (HPI) coils (placed at three fiducial points: nasion, left and right pre-auricular points). By periodically energising these coils the head position within the MEG sensor array was identified. Prior to data acquisition, the HPI coil locations and the subject's head shape were digitised using a Polhemus Isotrack system. Structural MR images for each subject were acquired using a Philips Achieva 3T MRI system (MPRAGE; 1mm isotropic resolution, 256x256x160 matrix, TR=8.1ms, TE=3.7ms, TI=960ms, shot interval=3s, flip angle=8 degrees and SENSE factor 2). We then converted the data into **SPM8** format (*FIL, UCL*) for further analysis. Two subjects were discarded after acquisition.

5.5.1.2 *APOE cohort*

Participants: 48 healthy volunteers were recruited (including 25 males) with a mean age of 47 and standard error of 0.13 years. Two subjects were discarded at the stage of data acquisition. The remaining cohort comprised three sub-groups, based on APOE genotype: $\epsilon 2$ carriers - 12 participants (5 males); $\epsilon 3$ homozygotes - 16 participants (9 males); $\epsilon 4$ carriers - 18 participants (11 males).

Data acquisition: Each participant was scanned (in a seated position) according to the following paradigm:

- **Session 1**: 6 minutes of eyes-open resting-state with a fixation cross.
- **Session 2**: 12 minute memory task - encoding phase.
- **Session 3**: 6 minutes of eyes-open resting-state with a fixation cross.
- **Session 4**: 12 minute Stroop counting task.
- **Session 5**: 12 minute visual oddball task.

- **Session 6**: 12 minute memory task - retrieval phase.

In this chapter, we discarded the four task sessions and focused exclusively on the two resting-state sessions. The MEG data were acquired using an *Elekta Neuromag* system. The data were acquired at a sampling frequency of 1kHz. Head localisation within the MEG helmet was achieved using four electromagnetic head position indicator (HPI) coils (two placed behind the ears, two placed on the forehead). The HPI coils were continuously energised to allow continuous tracking of the head position. Prior to data acquisition, the HPI coil locations and the subject's head shape were digitised using a Polhemus Isotrack system. Structural MR images for each subject were acquired at the University of Oxford Centre for Clinical Magnetic Resonance Research (OCMR) using a 3-T Siemens Trio scanner (*Siemens AG, Erlangen, Germany*) with a 12-channel head coil. The neuroimaging protocol included a 3D T1-weighted structural scan (MPRAGE, repetition time (TR) = 2040 ms, echo time (TE) = 4.7 ms, inversion time (TI) = 900ms, flip angle = 8 degrees, field of view = 192 mm, voxel dimension = 1 mm isotropic, scan time = 12 min 02 sec). During the MEG scan, eye movements were recorded electrically with vertically and horizontally orientated electrooculograms and optically with an EyeLink Eye Tracker (*SR Research Ltd., Mississauga, Ontario, Canada*). The electrocardiogram was also recorded.

The MEG data were down-sampled to 250Hz and channels with large amplitude, short duration artefacts identified (refer to section 3.2 for a more detailed discussion of the impact of channel artefacts on $MaxFilter^{TM}$). We then applied signal-space-separation to each session, via $MaxFilter^{TM}$ [50, 127]. We also used $MaxFilter^{TM}$ to apply movement compensation to the MEG data. We converted the data into **SPM8** format for further analysis. At this stage we discarded two additional subjects due to poor quality data.

5.5.1.3 *Simulated cohort*

We also developed a simulated data set for further validation of the MEG-adapted dual regression framework and to better understand the underlying phenomena that could drive a detectable group difference.

We started by generating a set of 25 spatial maps for the group. One map consisted of 2 crosses, corresponding to a 2-node network (see Figure 5.4A). This network had specific sub-group differences imposed. The remaining maps consisted of randomly located pairs of spherical nodes. These maps were kept constant in all the simulations. For each session's set of spatial maps, we added Gaussian noise.

For each map/session, an associated independent time course was simulated from a random [Laplace + Gaussian] distribution. This mixed distribution was used as it matched the leptokurtic distributions of the independent time courses estimated from real MEG data. The maps and time courses were multiplied and the resulting simulated true envelopes, $\mathbf{E}_n$, were imposed on randomly generated carrier data to produce the simulated true neuronal activity, $\mathbf{X}_n$. Gaussian noise was added to $\mathbf{X}_n$. The simulated data were bandpass filtered into the 8-13Hz range. The band-limited data were treated as being equivalent to the output of the beamformer used on real data, $\hat{\mathbf{X}}_n$, and were analysed similarly. Note that in this simulation framework, we did not model the projection of sensor-level noise and subsequent *weights normalisation*: consequently we made the simplifying assumption that the *weights normalisation* at each voxel was unity (i.e. $\tilde{\mathbf{X}}_n = \hat{\mathbf{X}}_n$).

We performed 8 simulations:

1. No group difference.

2. Difference in both the mean and the variance of the component time course.

3. Difference in the variance of the component time course.

4. Difference in the mean of the component time course.

5. Difference in the position of one of the nodes of the spatial map.

6. Difference in the amplitude of one of the nodes of the spatial map.

7. Difference in the amplitude of the spatial map outside of the two main nodes.

8. Global difference in the signal-to-noise ratio of the data.

Figure 5.4A shows simplified diagrams of the 8 simulations. These simulations were designed to assess the sorts of differences that can be detected in our analysis and not to be a full simulation of MEG data.

5.5.2 Preprocessing

We applied the standard preprocessing pipeline (described in detail in section 3.6) to the *eyes-closed/active-state* and *APOE* data sets. For each session, we used independent component analysis at the sensor level to remove physiological and environmental artefacts. We then visually inspected each session of data. We identified any channels or periods of data that looked corrupted and flagged them as bad. At this stage the *eyes-closed/active-state* data was resampled from 600Hz to 200Hz. Each subject was co-registered to the MNI template brain by mapping the subject's head shape (measured with the Polhemus Isotrack) to their structural and to the MNI canonical brain using **SPM8** (*FIL, UCL, London*).

Every session was band-pass filtered into a single frequency band. For the *eyes-closed/active-state* cohort, the data were filtered into the alpha band (8-13Hz). We chose the alpha band as we wished to validate our analysis on the well-characterised increase in visual cortex alpha power in *eyes-closed* rest compared with *eyes-open* activity [20,33]. For the *APOE* cohort, the data were filtered into the 4-30Hz range as we did not have a specific frequency band to target.

Source reconstruction of each session was performed using a linearly-constrained minimum variance beamformer, described in chapter 3 [134, 140]. The lead-fields were estimated using the overlapping local spheres forward model [75] implemented in **FieldTrip** (*Donders Institute for Brain, Cognition and Behaviour, Radboud University*). The beamformer was used to estimate source space activity across a 6mm grid spanning the whole brain.

We estimated the windowed-average oscillatory amplitude envelope using 2s non-overlapping windows. A 2s window was selected based on the optimisation analysis presented in section 4.3 [88]. We estimated the *weights-normalised* and *non-weights-normalised* envelopes for use in stages 1 and 2 of the MADR analysis. We spatially

smoothed each session's envelopes with a 4mm Gaussian kernel using **FSL** (*FMRIB, Oxford*). We spatially re-sampled the data to an 8mm grid. This was applied to both the *weights-normalised* and *non-weights-normalised* envelopes in parallel.

5.5.3 Application of MEG-adapted dual regression

***Eyes-closed/active-state cohort*:** For the *eyes-closed/active-state* data, we validated MADR by using it to detect alpha power increases in the visual networks during *eyes-closed* rest. We used a set of 18 RSN maps derived from a spatial ICA performed on fMRI data [122]. Using this fMRI basis set, we performed a paired t-test between each subject's *eyes-closed* and *active-state* sessions. We applied the test in three ways for each component: 1.) at every voxel in the set of session-specific spatial maps; 2.) for the mean values of the component time course; 3.) for the variances of the component time course. We applied spatial smoothing with a 16mm Gaussian kernel to the variances of the regression parameter estimates for the voxel-wise paired t-test. Spatial smoothing was critical to overcome the poor estimate in the variance of the residual at the group level. Spatial smoothing assumes that the expected value of the variance of the residual is approximately uniform across the brain. However, because our first level regression parameters were estimated from the *non-weights-normalised* envelopes, they have an imposed depth bias that is reflected in the variance of the parameter estimates at the group level. In order to account for this, we divided every subject-level regression parameter map by the group mean *weights normalisation* map.

We performed the MADR analysis on the *eyes-closed/active-state* data a second time but deliberately used the *weights-normalised* envelopes throughout. We did this to demonstrate the effect of using *weights-normalised* data to infer differences across sessions in the context of MADR.

***APOE cohort*:** For the *APOE* cohort, we used a MEG-derived spatial basis set. Specifically, we used a set of 25 spatial maps that we estimated by performing a temporal ICA decomposition on the concatenated windowed-average envelopes using

the analysis described in section 4.5.3. These maps are the electrophysiological analogue of the RSN maps estimated from spatial ICA on fMRI BOLD data that we used in the *eyes-closed/active-state* analysis. In order to estimate the MEG-derived functional networks, we de-meaned each voxel of the *weights-normalised* envelopes and then temporally concatenated all the sessions. We normalised this concatenated matrix by the global standard deviation. We then used ICA to decompose the concatenated envelopes into 25 temporally independent time courses (implemented using **FastICA**) [77]. This yielded a set of independent time courses and spatial maps for the concatenated group.

We used the spatial maps as the basis set for the MADR analysis to get 25 spatial maps for each session of data. Using these session-specific maps, we performed a single group analysis, contrasting the $\epsilon 3$ homozygote group with the $\epsilon 4$ carrier group. However, as we had two sessions of resting-state data, we performed an intermediate subject-level fixed effects analysis to collapse the two session-level regression parameters/component means/component variances into subject-level estimates. As with the *eyes-closed/active-state* analysis, we performed 3 sets of tests: on the spatial maps; on the means of the component time courses; on the variances of the component time courses. For the voxel-wise statistics, we performed smoothing on the variance of the residuals of the group level GLM using a 40mm Gaussian kernel. In order to account for the depth bias in the variance of the group-level residuals, we divided every subject-level regression parameter map by the group mean *weights-normalisation* map.

***Simulated cohort*:** When analysing the simulated data, we performed an identical MADR analysis using a paired t-test to the *eye-closed/active-state* data.

5.5.4 Results

5.5.4.1 *Simulated data*

Figure 5.4 shows the results of the MADR analysis on the eight simulations. Figure 5.4B shows the resulting Z-statistics from the paired t-tests on the component means

and variances. As we knew *a priori* which component to test, we only had to perform tests on one component for each simulation and as such did not need to correct for multiple comparisons over components. We defined significance as $p \leq 0.01$. Figure 5.4C shows t-statistical maps for the voxel-wise paired t-test performed across the session specific spatial maps. We used threshold-free cluster enhanced (TFCE) permutation testing (implemented in **FSL**'s **RANDOMISE** (*FMRIB, Oxford*)) to account for multiple comparisons over voxels. Any significant clusters have been shaded in green.

1. **No Group Difference**: the dual regression has not found any significant false positive group differences.

2. **Differences in the mean and variance of component time courses**: the statistics on the mean and variance have both accurately detected a significant group difference. The voxel-wise GLM has found a significant difference across the whole simulated network.

3. **Differences in the variance of component time courses**: the analysis has correctly detected a group difference in the variance of the component time course between sessions. It has also correctly detected a significant difference across the majority of the network from the voxel-wise GLM. However, it has falsely inferred a difference in the group means.

4. **Differences in the mean of component time courses**: a significant difference has been found in the means of the component time course but not in the variance. No significant differences were found in the voxel-wise analysis of the spatial maps.

5. **Difference in the position of one node of the spatial map**: the dual regression technique has performed perfectly here: it has detected no group differences in the component mean and variance but has correctly detected the spatial difference due to the shifted network node in the voxel-wise analysis.

6. **Difference in amplitude of one node of the spatial map**: the dual regression framework has inferred differences in the mean and variance of the component time courses that were not directly simulated. It has correctly detected a significant negative group difference between the two simulated groups, exactly matching the node with reduced amplitude. However, it has incorrectly inferred additional weaker but still significant differences in the second network node.

7. **Difference in sub-threshold amplitude outside of the main spatial map**: In this simulation variant (that is very similar but weaker than simulation 6) the analysis has correctly inferred no group differences in the mean and variance of the component time courses. It has found significant differences in the correct sub-threshold region where the group difference was simulated but it has not accurately delineated this region.

8. **Difference in the signal-to-noise ratio**: The analysis has detected significant differences in the component means and variances. It has also found a small significant cluster of voxels showing a group difference in the voxel-wise analysis.

5.5.4.2 *Eyes-closed/active-state data*

We performed an MADR analysis on the *eyes-closed/active-state* data using a fMRI spatial basis. Specifically, we used 18 spatial maps estimated from spatial ICA performed on resting-state fMRI data [122]. For each component, we estimated a component mean, variance and spatial map specific to each session of MEG data. Considering the variances of the components, we found 5 components that showed significant differences in component variance between *eyes-closed* and *active-state* sessions. We discarded the other 13 components. Figure 5.5 shows the results of the dual regression for 3 of the significant components: the lateral visual network; the visual network; the bilateral hippocampi. For each component, the Z-statistical map of the spatial ICA group result is shown in red/yellow. Next to each spatial map is the statistical

map of unity minus the corrected p-values. All three components have significant increases in oscillatory alpha power variance in *eyes-closed* sessions relative to *active-state* sessions. For the lateral visual network, this is confined to the left hemisphere; for the visual network this difference is in both the left and right medial visual areas; for the hippocampal component, we found significant differences confined to the left temporal lobe. The hippocampal component also had a significant elevation in the mean of the component time courses in the *eyes-closed* sessions relative to the *active-state* sessions. The cerebellum and bilateral primary somatosensory cortices (PSSC) components also showed significant increases in their component mean and variances. For the cerebellum: $p_{\text{corrected}}$ of mean $= 0.01$; $p_{\text{corrected}}$ of variance < 0.01. The voxel-wise GLM confined these differences to the cerebellum. For the PSSC component: $p_{\text{corrected}}$ of mean $= 0.06$; $p_{\text{corrected}}$ of variance < 0.01. The voxel-wise GLM did not find any significant clusters.

In order to emphasise how MEG-adapted dual regression should not be performed using *weights-normalised* envelopes, we performed an identical analysis to that shown in Figure 5.5 but using the *weights-normalised* envelopes throughout the dual regression. We found no significant components showing *increased* alpha activity in *eyes-closed* versus *active-state*. Instead, we found two components, corresponding to the medial visual network and caudate, that showed significant *decreases* in alpha power variance in the voxel-wise analysis of the session-specific maps. This result is shown in Figure 5.6. Note that we used a green-purple colour scale to denote the unity minus the p-value for the *active-state* - *eyes-closed* rest contrast. In agreement with the voxel-level analysis in Figure 5.2, using the *weights-normalised* envelopes to estimate differences across sessions actually reversed the effect direction.

5.5.4.3 *APOE data*

We applied the MEG adapted dual regression analysis to the output of the temporal ICA performed on the 4-30Hz envelopes of the *APOE* data set. This yielded a statistical map for each of the 25 components, as well as a statistic for the mean and variance. We focused on the contrast between homozygous $\epsilon 3$ carriers versus $\epsilon 4$

carriers. Figure 5.7 shows the results for the right and left frontoparietal networks. Both RSNs show increased variance in oscillatory power in the homozygous $\epsilon 3$ group compared to the $\epsilon 4$-carriers.

5.6 Discussion of MEG-adapted dual regression

5.6.1 Interpretation of dual regression

The simulations were very powerful in simultaneously highlighting the strengths and weakness of the MEG-adapted dual regression approach. Firstly and most importantly, the procedure is robust against false positives. When no group differences were simulated, the only significant differences were found in a few voxels in the voxel-wise GLM analysis. However, when multiple comparisons over voxels were taken into account, these voxels did not pass significance. We conclude that false positives should only be an issue when multiple comparisons are not being addressed correctly.

A weakness in the dual regression approach is interpretability. It is extremely challenging to infer the underlying mechanisms from a given positive finding. Ideally, we would like to infer differences in functional connectivity from this analysis. Unfortunately, this is not trivial. MEG-adapted dual regression can accurately localise where informative group differences exist. It can also attribute these group differences to specific resting-state networks. However, other complementary analyses would be needed to elucidate the physiological mechanisms that create these group differences.

5.6.2 Mean, variance and voxel-wise statistics

We have developed three related statistics that can be estimated for each component: mean of the time course, variance of the time course/spatial map and voxel-level statistics. These three statistics offer overlapping insights into the underlying group differences. Our suite of simulations helped to disentangle these insights.

The mean of the component (for a given session) corresponds to an estimate of the average oscillatory power associated with that component for that session. When we

simulated a difference in the mean value associated with a component between two groups, we accurately inferred this difference. However, when we simulated a difference in the variances of the component between two groups, we incorrectly inferred a difference in the mean as well as a difference in the variance. This suggests that our estimates of mean and variance for each session and component are not independent.

This intuitively makes sense as differences in session-specific component variances will lead to different session-specific maps (because the variance normalisation in MADR stage 2 shifts any component time course variance into the spatial maps). These different session-specific spatial maps will be used in equation 5.13, resulting in the estimation of different session-specific means.

Our conclusion at this stage is that inferred differences in component-level means most likely correspond to true differences if no difference in variance is detected. If both mean and variance differences are detected then we cannot demonstrate at this stage whether there is a genuine change in the mean oscillatory power or just a change in the variance of oscillatory power, leading to an incorrect detection of a difference in the mean. When considering the variances of a given component between sessions, we can be more confident as the dual regression analysis maps variance correctly between component- and voxel-levels.

As stated in the exposition of the MADR method, the variance of the component is estimated as the overall variance of the spatial map for a single session. One might argue that we gain nothing by estimating the component-level variance as it is implicit in our voxel-wise analysis. However, utilising the component-level variance has two advantages. Firstly, we can avoid performing statistical tests over voxels, thus decreasing the severity of the multiple comparison correction. Secondly, when we performed voxel-level statistical tests, we found that it was important to spatially smooth the variance of our estimates of the residuals at each voxel. In order to allow this, we had to impose a group-level *weights normalisation*, using the group average *weights normalisation* map. This is an approximation that is not ideal as the *weights normalisation* should be specific to each session. One alternative solution might be to apply anisotropic variance smoothing to each session along the contours of constant

weights normalisation specific to each session. Performing component-level statistics on the variance completely bypasses this issue.

In this study, we used the component-level variances to apply a weak control of the family-wise error rate of multiple comparisons over components. Ideally, one would develop a single, unified framework that controls for multiple comparisons over voxels and components but that is beyond the scope of this work. Despite the additional challenges and computational cost of performing voxel-wise statistics, we argue that they are more informative and sensitive. As shown by simulation 7, when we simulate a localised, small group difference in a sub-threshold region of the spatial map, only the voxel-wise analysis is sensitive enough to detect this group difference. Essentially, there is more that is the same between the two groups than there is that is different. Hence, component-level analyses cannot detect those subtle differences.

5.6.3 Validation of MADR - inferring differences between *eyes-closed* rest and *active-state*.

We validated the MADR framework by using it to infer differences in resting-state networks during *eyes-closed* rest versus *eyes-open* visual stimulation via a movie. We expected strong alpha power increases in the visual cortex (supported by our voxel-level analysis in Figure 5.2). Using RSN maps derived from spatial ICA performed on fMRI data, we found increased variance in the 8-13Hz band-limited power associated with the lateral visual network and the visual network. This aligns strongly with the well-established increased visual alpha power observed in *eyes-closed* rest [20, 33].

5.6.4 Novel insights of MADR - differences in resting-state activity between $\epsilon 3$ homozygotes and $\epsilon 4$ carriers.

Analysis of resting-state functional connectivity shows promise in helping elucidate the mechanisms driving Alzheimer's disease. The default mode network (DMN) has been shown to have deficient activity in the posterior cingulate cortex (PCC) [64]. Further evidence that implicates abnormal DMN activity to AD pathology is the

strong correlation between amyloid deposition and the DMN [32]. Much work has been done to investigate if the APOE-ϵ4 genotype has an effect on resting-state brain function, with a particular focus on DMN. In fMRI, Filippini et al. showed, using spatial ICA and dual regression, that ϵ4 carriers have abnormal activity in their DMNs compared to the more common homozygous ϵ3-carrier population [55]. Furthermore, Trachtenberg et al. found additional differences in the auditory network, lateral visual network and left frontoparietal network using similar methods [129].

Our preliminary exploration of the effect of APOE genotype on resting-state brain function showed strong elevations in the variance of 4-30Hz oscillatory power associated with both the right and left frontoparietal networks. This is in agreement with previous fMRI findings by Trachtenberg et al. (2012) who detected APOE-related changes in the left FPN [129]. Their differences were detected in the caudate, thalamus, anterior cingulate cortex and parietal cortex. We also detected differences in the caudate but the dominant group differences were in the main nodes of the FPNs.

At this stage, it is not clear why both studies find differences associated with the left FPN network but in MEG the differences include the main nodes of the network, whilst in fMRI they are only external to the main network. There is no obvious justification for fMRI being insensitive to within-network differences. However, we found in our development of MEG-adapted dual regression that spatial smoothing of the variance of the regression parameter estimates was critical to improving our sensitivity to within network differences. Trachtenberg et al. do not state that they perform any equivalent smoothing and this may contribute to the difference between the two studies. Of course, the disagreement may be due to the difference in what the two modalities are measuring: the MEG analysis is driven by band-limited power fluctuations of neural oscillations while the fMRI is measuring a vascular response to increased metabolic demand of neuronal populations.

Our ICA decomposition of the band-limited power time series did not produce a single default mode network and so we were unable to reproduce the DMN-specific findings of Filippini et al. and Trachtenberg et al. [55, 129]. However, this highlights one of the major opportunities afforded by the MADR framework: in future studies,

we can use the fMRI-derived ICA maps from previous studies as the spatial basis set to investigate directly if the differences observed in the BOLD signal RSNs are also reflected in the MEG data.[2]

5.7 Conclusion

In this study, we have laid the foundations for performing a multitude of group statistical analyses on resting-state or task-positive MEG data. Analysis of networks of spontaneous brain activity in the resting-state has proven to be a fruitful research area in fMRI neuroimaging. MEG can provide a unique and novel insight into this arena, particularly as it is not limited by the haemodynamic response function and has extremely rich temporal and spectral resolution. The framework that we have presented will be particularly useful in bridging the gap between the wealth of information already uncovered with fMRI and the relatively unexplored terrain of electrophysiological data in MEG.

We have identified *weights normalisation* as a potential major source of error if not accounted for properly. While our theoretical and practical demonstrations have been limited to beamformer-based source reconstructions, we have yet to establish whether other source reconstruction techniques (such as minimum norm estimates) are vulnerable to similar confounds. Minimum norm approaches have the potential to suffer from similar issues but we have not demonstrated the severity of the *weights normalisation* confound in such a context.

Finally, as with any exploration of functional connectivity, we must caution against the temptation of over-zealous interpretation. One of the key findings of this chapter is that many underlying processes can manifest as similar true positive findings in dual regression. Therefore, this framework alone is insufficient to draw conclusions about changes in functional connectivity in neural activity.

[2]Refer to the D.Phil thesis of V. Heise (*under examination*) where fMRI-derived RSN maps from the APOE cohort were used as a spatial basis set for MEG-adapted dual regression [71]

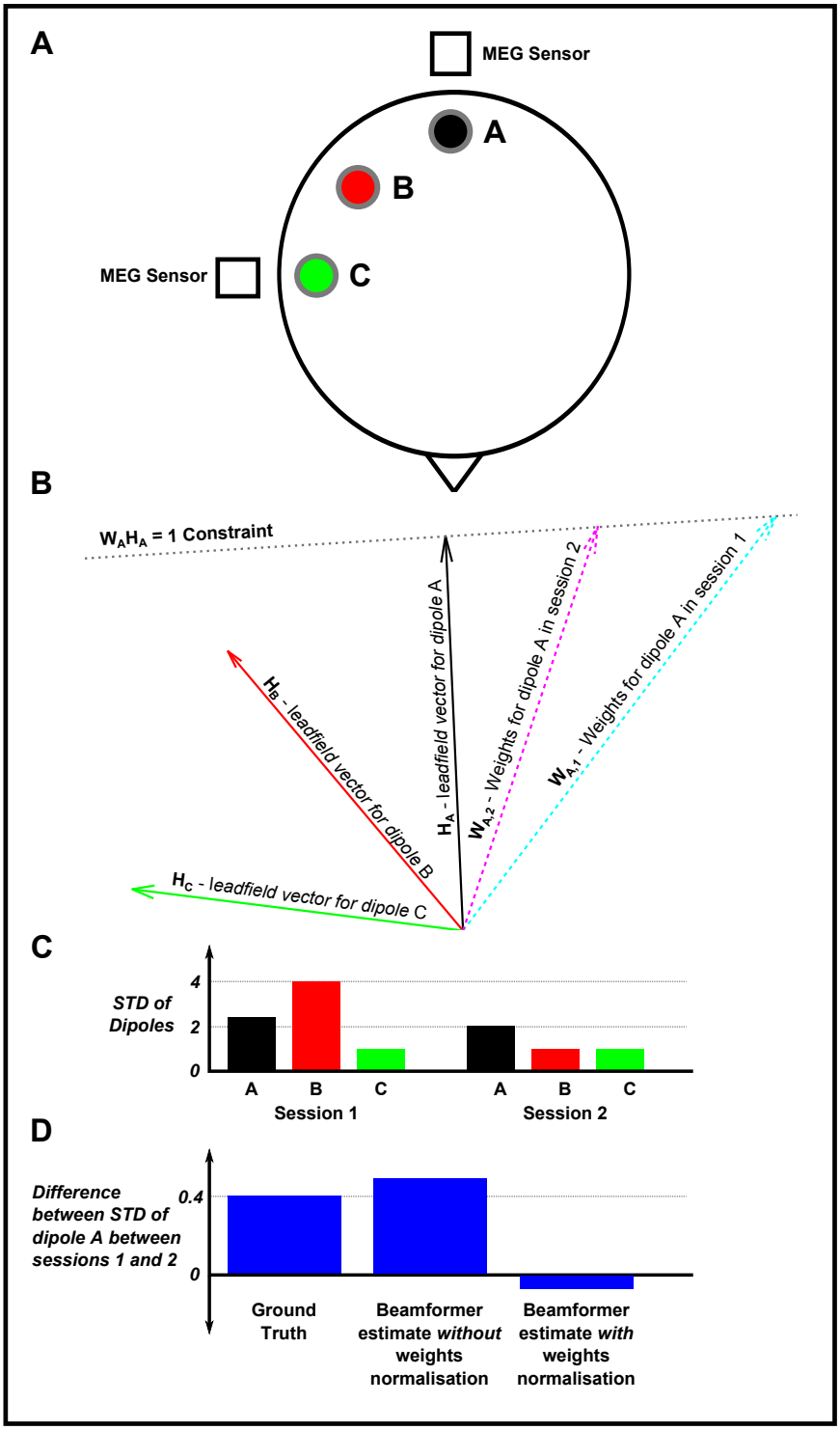


Figure 5.1: *A simulation showing one example of how weights normalisation can reverse the apparent direction of an effect. In this simulation, the effect we measure is the standard deviation of dipole A, σ_A, between two sessions (1 & 2) where $\sigma_{A,1} > \sigma_{A,2}$. A. The schematic of our 2-dimensional simulation, consisting of 3 dipoles and 2 MEG sensors. The dipole time courses are uncorrelated and normally distributed. B. The lead-field vectors for our three dipoles (solid black, red and green arrows) and the beamformer weights vectors for dipole A from session 1 and session 2 (dashed cyan and magenta arrows). C. The standard deviations of the three dipoles in sessions 1 and 2. D. The ground truth and beamformer estimated differences (**without** and **with** weights normalisation) between the standard deviation of dipole A in sessions 1 and 2.*

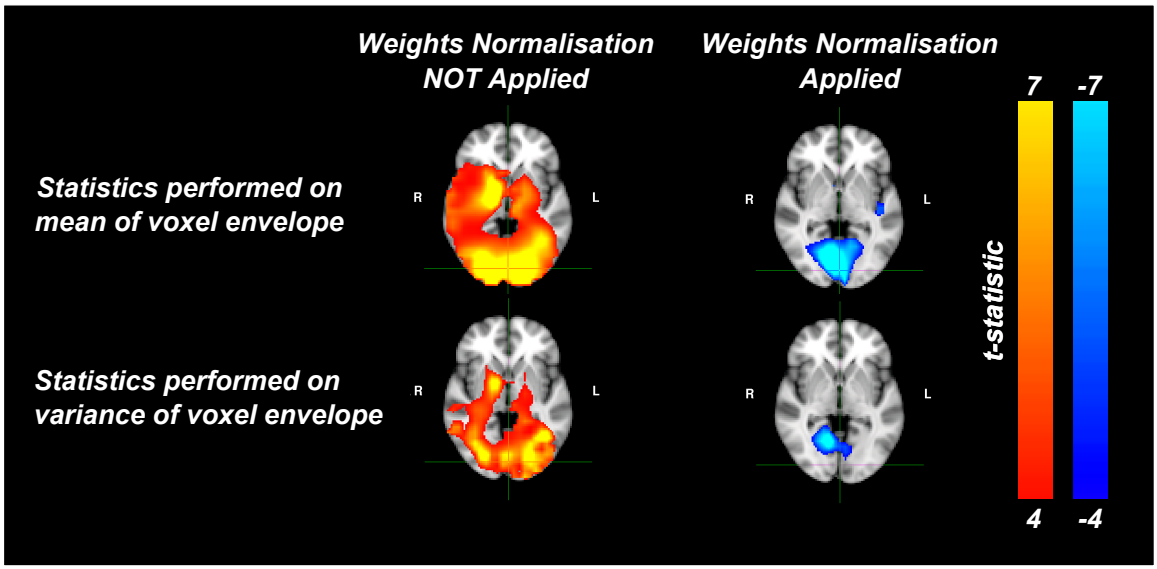


Figure 5.2: *Axial slices through the visual cortex showing results of a voxel-wise paired t-test performed on the mean and variance of the envelope of the alpha (8-13Hz) oscillations of the eyes-closed and active-state sessions before and after applying weights-normalisation. For each map, t-statistics were thresholded at 4, with positive values shown in red/yellow and negative values shown in blue. Without weights normalisation, we correctly measured a significant increase in the mean and variance of alpha power in the eyes-closed condition compared with the active-state condition. With weights normalisation, we detected a decrease in the mean and variance of alpha power in the eyes-closed condition compared with the active-state condition, demonstrating that the weights normalisation confound can actually reverse the underlying effect direction.* ***FSL's RANDOMISE*** *was used to perform threshold-free cluster enhanced (TFCE) permutation testing to account for multiple comparisons. All t-statistics shown are members of significant ($p_{corrected} < 0.05$) clusters.*

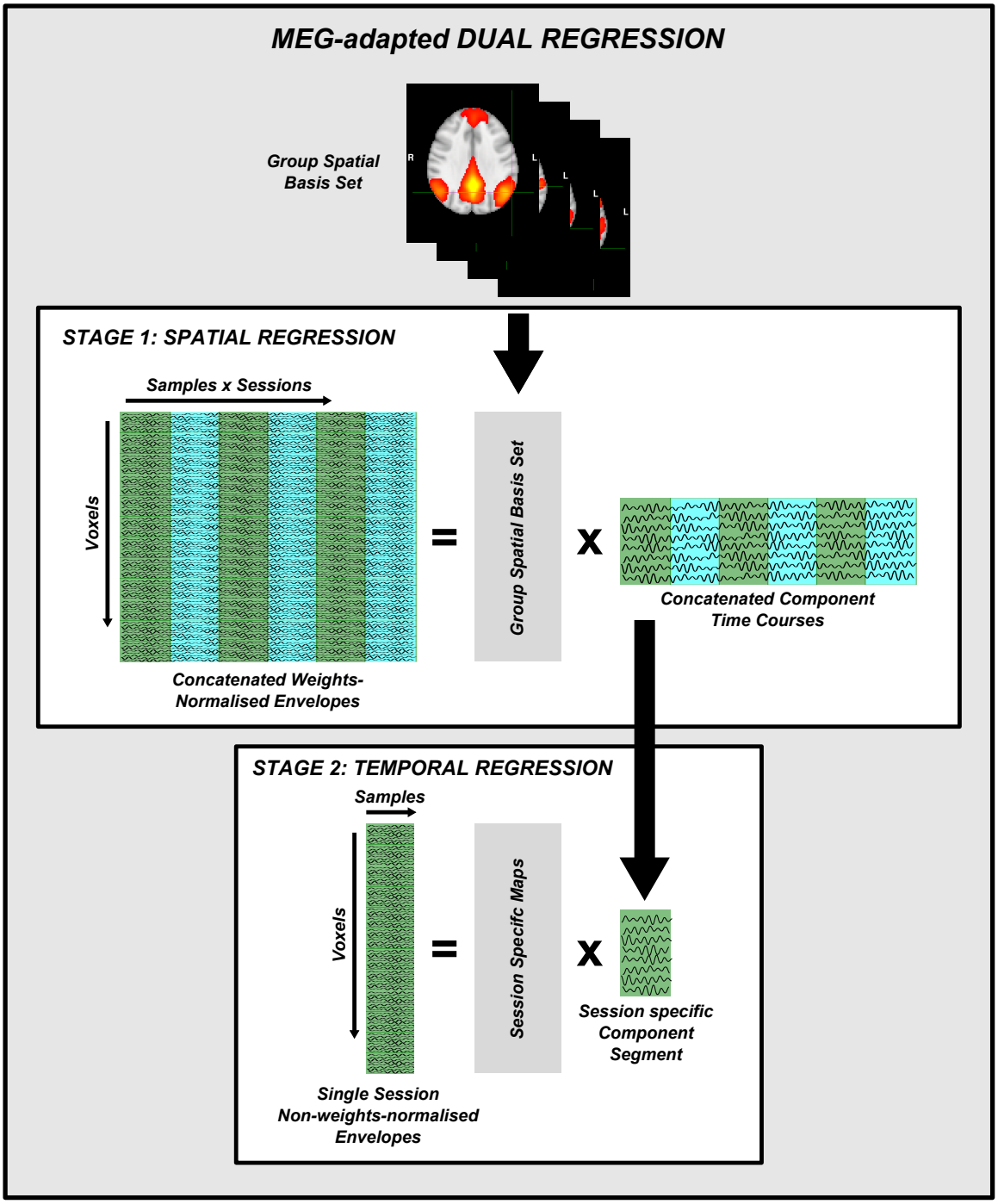


Figure 5.3: *A schematic of the MEG-adapted dual regression (MADR) analysis. We provide the MADR analysis with a spatial basis set as an input.* ***Stage 1*** *of MADR consists of a spatial multiple regression of the basis set from the group concatenated, weights-normalised envelopes to give group-level component time courses which are then broken up into session-specific blocks.* ***Stage 2*** *consists of a temporal multiple regression of each session-specific block from the non-weights-normalised envelopes to give session-specific spatial maps.*

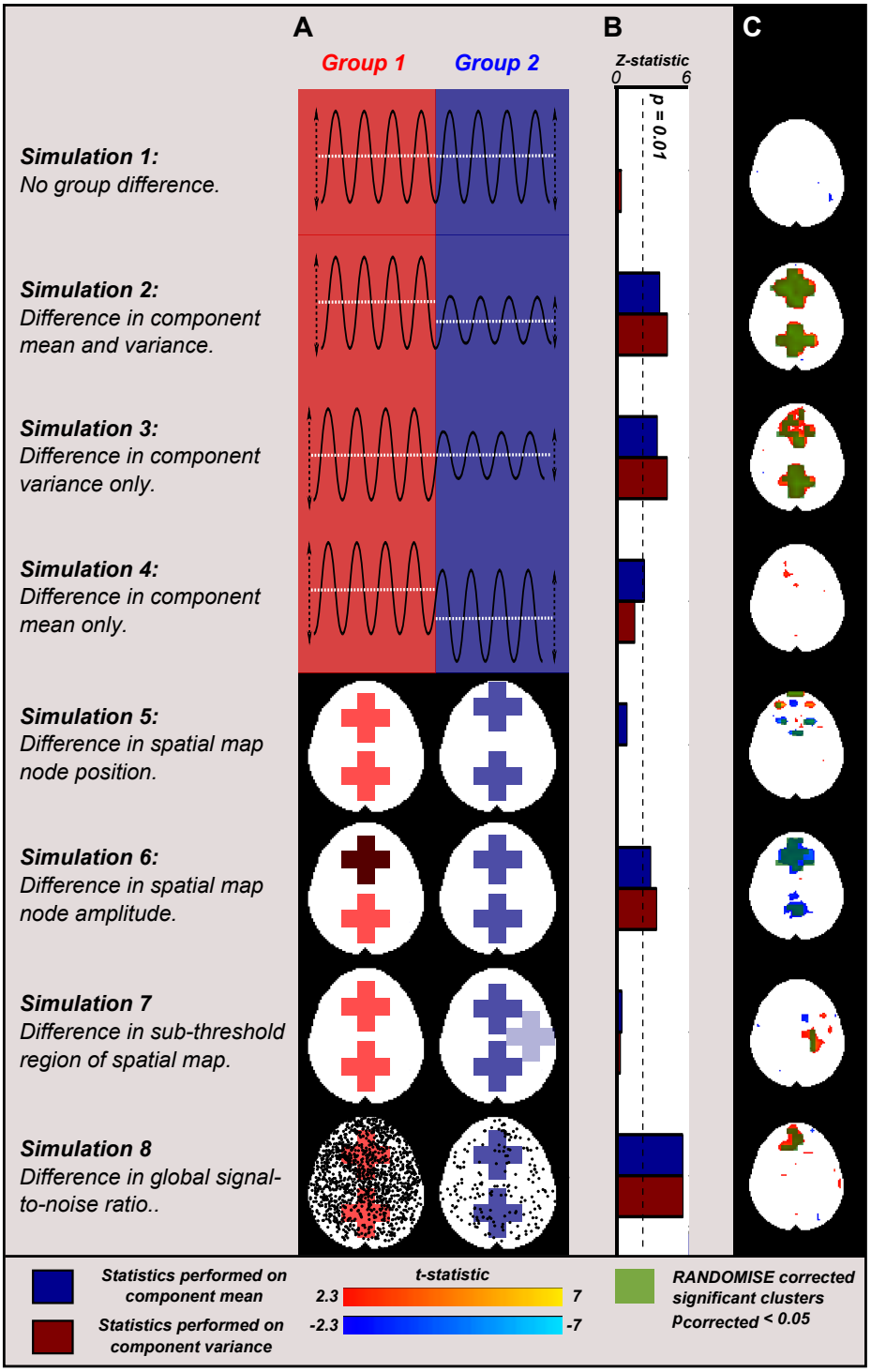


Figure 5.4: *Simulations to demonstrate the performance of the MEG-adapted dual regression analysis. A. Schematics of the 8 simulations used. Simulation 1 tests performance under no group difference. 3 simulations (2, 3 and 4) impose group differences in the mean and/or variance of the component time course. 3 simulations (5, 6 and 7) impose group differences in the spatial map of the component. Simulation 8 shows the effect of differences in global signal-to-noise ratio. B. Results of the component-level statistics. The paired t-test was applied to the estimated component means (blue bars) and variances (red bars). C. The spatial maps were subjected to an equivalent paired t-test at each voxel: positive t-statistics are shown in red/yellow; negative t-statistics are shown in blue; significant clusters (found with threshold-free cluster enhanced (TFCE) permutation methods using* ***FSL****'s* ***RANDOMISE****) are shown in green.*

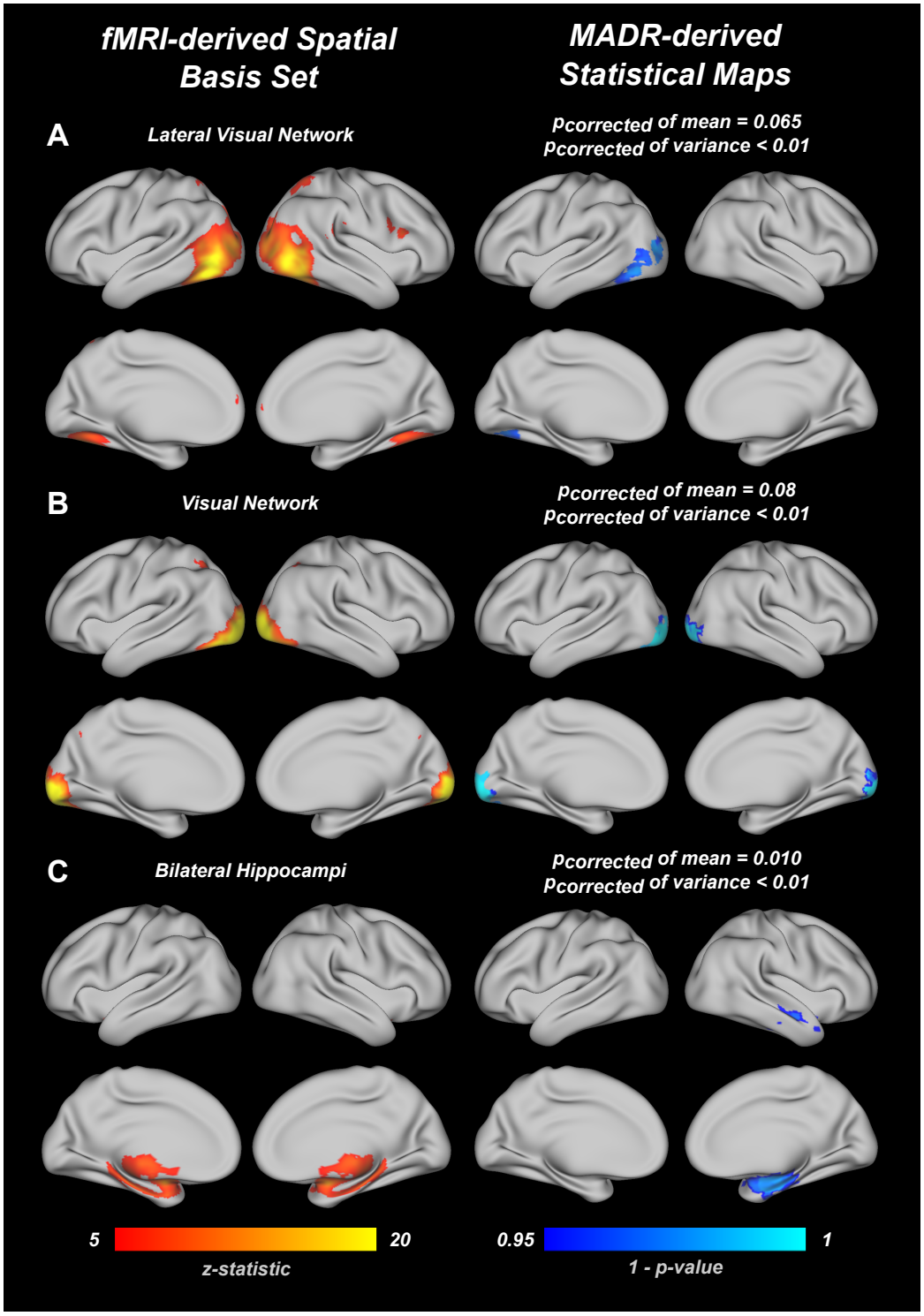


Figure 5.5: *Results of the MADR analysis on eyes-closed/active-state data using fMRI spatial ICA maps as a basis set [122]. Three fMRI-derived RSN components are shown: A. lateral visual network; B. visual network; C. bilateral hippocampi. In Red/Yellow (left): spatial ICA Z-statistical maps, thresholded at $Z = 5$. In Blue (right): maps of unity minus the p-values for the voxel-wise paired t-test between eyes-closed and active-state conditions, thresholded at 1-$p_{corrected} = 0.95$. P-values were corrected for voxel-wise multiple comparisons using **FSL**'s **RANDOMISE** threshold-free cluster enhanced (TFCE) permutation methods. For each component, the corrected p-value for the paired t-test applied to the component time course means and variances are also shown. Component p-values were corrected using a multi-step up test with significance threshold of $p \leq 0.01$ [101].*

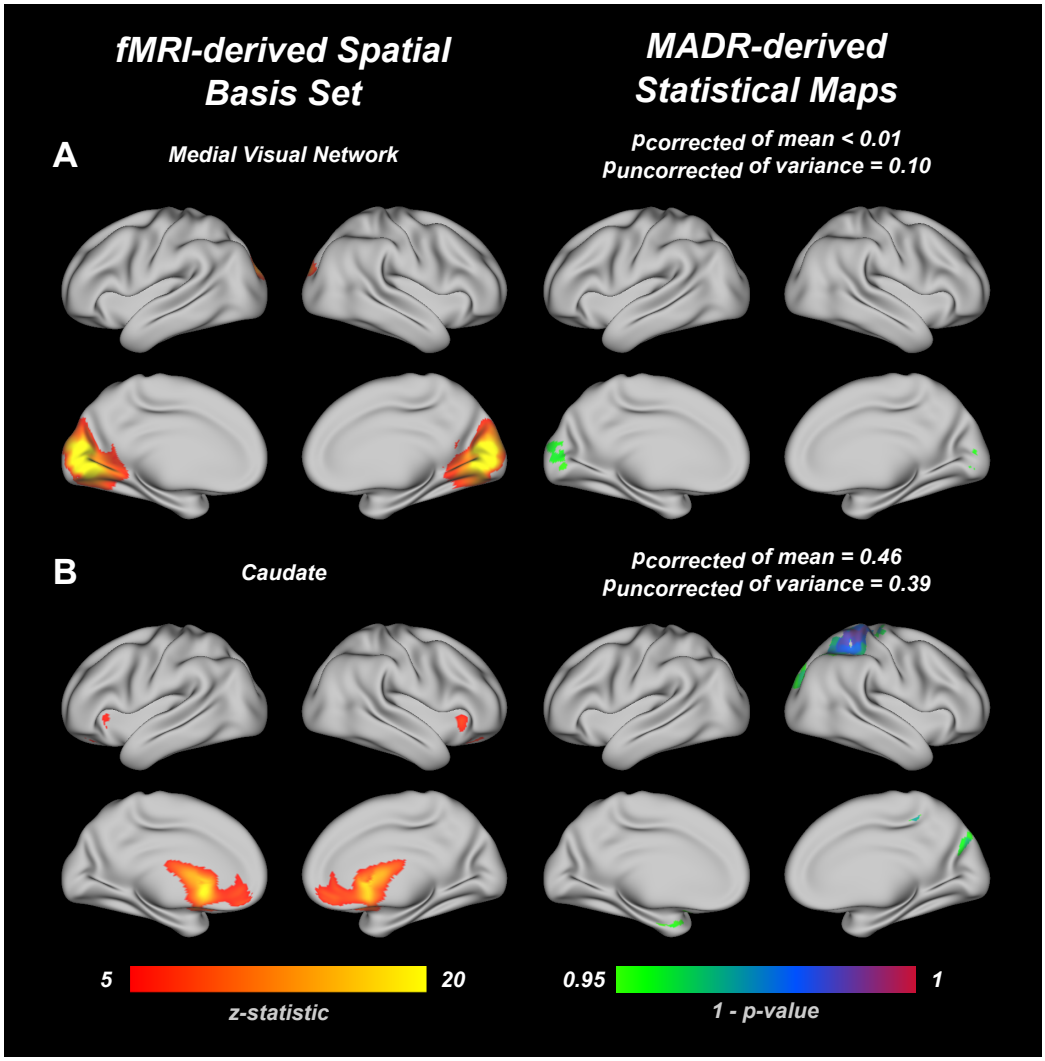


Figure 5.6: *Results of the MADR analysis on eyes-closed/active-state data using fMRI spatial ICA maps as a basis set,* ***but when incorrectly using the weights-normalised data throughout the analysis****. No components showed significant clusters for the eyes-closed - active-state contrast. Two components (A. medial visual network and B. caudate) showed significant clusters for the active-state - eyes-closed contrast (for which unity minus p-values are shown in green/purple), demonstrating that weights normalisation can actually reverse the underlying effect direction in an MADR analysis.*

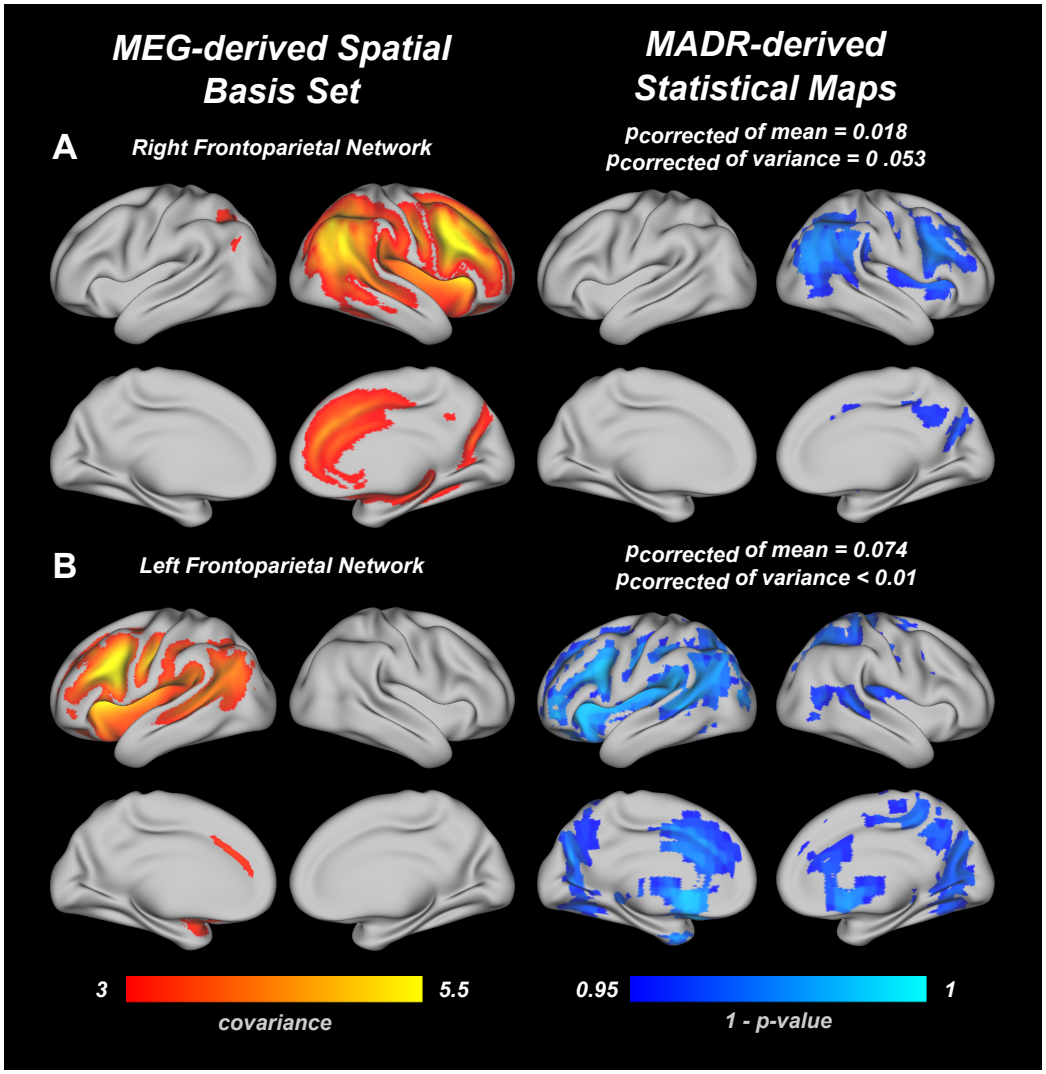


Figure 5.7: *Results of the MADR analysis on the APOE data using an MEG-derived spatial basis set. Two components are shown: A. right frontoparietal network and B. left frontoparietal network. In Red/Yellow: maps of covariance between the independent time course of the component and the concatenated envelopes at each voxel. Covariance maps have been normalised to have unit variance. In Blue: maps of unity minus the p-values for the group-level GLM testing for increases in the $\epsilon 3$ homozygotes relative to $\epsilon 4$ carriers. P-values were corrected for voxel-wise multiple comparisons using* ***FSL****'s* ***RANDOMISE*** *threshold-free cluster enhanced (TFCE) permutation methods. For each component, the corrected p-value for the GLMs applied to the component time course means and variances are also shown (again for the $\epsilon 3$ homozygotes - $\epsilon 4$ carriers contrast). P-values were corrected using a multi-step up test with significance threshold of $p<0.05$.*

Chapter 6

Analysis of functional networks in task-positive paradigms

6.1 Chapter abstract[1]

In this chapter, we demonstrate that our functional connectivity analysis, based on performing temporal ICA on the slow-moving envelopes of beamformer-estimates of band-limited neural activity, can be used to extract functionally-connected networks from task-positive MEG data. We present a *post hoc* analysis of the independent components using a mixed-effects General Linear Model (GLM). We show that our combined ICA/GLM analysis can provide advantages in task-positive MEG analysis which are unavailable to either technique when used in isolation. In particular, it is a powerful tool for reducing the number of statistical tests, thus ameliorating the problem of multiple comparisons. We apply our analysis to a *2-back* working memory experiment and show that it matches previous features associated with working memory function, outperforms a standard voxel-wise GLM analysis, and provides novel insights into working memory function in MEG.

[1]The analysis in this chapter has been published in full by Luckhoo et al. in 2012 [88]. The descriptions of the data, methods and results have been taken from this publication.

6.2 Introduction

Throughout this thesis the focus has been on the identification and analysis of functionally-connected networks in the resting-state. In this chapter, we change tack to consider whether equivalent networks can be detected in task-positive MEG data and if so how we might analyse them. We have referred to these networks in previous chapters as resting-state networks (RSNs), in this chapter we will term them functionally-connected networks (FCNs) to reflect that these networks are found in task-positive data. Smith et al. (2009) showed, using a meta-analytic comparison of fMRI derived RSNs (extracted using group spatial ICA) with the BrainMap task activation database, that RSNs have a functional correspondence to task-related regions [122]. Ultimately, this idea motivates the field of research in resting-state networks: we are not only interested in understanding the resting brain in isolation, but also want to use it to further our understanding of general cognitive function, dysfunction and disease.

In this chapter, we apply our ICA framework (described in section 4.5.3) to a working memory paradigm. We show that functionally-connected networks are present during working memory processing that are similar to those found in the resting-state and that we can extract them in an identical fashion to resting-state analyses. We then proceed to develop a framework for applying mass-univariate statistics to infer which networks are actually involved in the task.

6.2.1 Coupling ICA and the GLM to ameliorate multiple comparisons and inform component classification

In chapter 5 we considered one interpretation of RSNs as a spatial basis set which can be used as a functionally relevant, low-dimensional representation of whole brain activity. In this chapter, we use the FCNs extracted from the task-positive data in much the same way. Statistical analysis of task-positive MEG data is nothing novel: mass-univariate General Linear Modelling (GLM) and pseudo-T-statistical analyses have been tailored to MEG data to find regions that exhibit task-modulated neural

activity [24,60,136]. However, in order to control the family-wise error rate such that task-dependent activations can be reliably inferred, one must correct for multiple comparisons of statistical tests [101, 139]. In MEG, multiple comparison corrections are particularly important, as statistical tests can be performed over time, space (either at the sensor-level or source-level), and frequency band.

Several techniques for controlling the family-wise error rate due to multiple comparisons have been adapted for MEG. Barnes and Hillebrand (2003) exploit the inherent spatial smoothness of beamformed MEG data to efficiently account for multiple comparisons [11]. Their technique uses the correlations between beamformer weights to map the inhomogeneous spatial smoothness of source-reconstructed MEG data, which is then used to statistically flatten the source-space statistical maps. Singh et al. (2003) demonstrate that non-parametric permutation techniques can be applied to account for multiple comparisons in source-level, multi-subject, MEG studies [120]. Permutation techniques estimate a data-derived null distribution of cluster sizes from which it is possible to infer the probability that a given statistical activation cluster could occur by chance [102]. Both statistical flattening and permutation techniques capture the dependence between neighbouring voxels due to spatial smoothness. Here, we use the functionally-connected networks as a representation of the brain that effectively captures the functional smoothness: i.e. we are clustering voxels based on functional similarity in their band limited power time courses. We can then perform statistics at the component-level, rather than at the voxel-level. We demonstrate that this is a powerful alternative solution for ameliorating the multiple comparisons associated with voxel-level statistics.

One unavoidable aspect of ICA is that the components are produced in a random order with no immediate way to classify which components are of interest (i.e. corresponding to functional networks) and which components are irrelevant (i.e. corresponding to non-neural artefacts or residual neural activity not associated with functional networks). In chapter 4 we identified the relevant RSN components by comparing them with fMRI-derived RSNs. In chapter 5 we identified the relevant components based on previous work in fMRI to consider the effect of the APOE-ϵ4

genotype on the frontoparietal networks. In both cases we used prior knowledge to determine which components to consider and which components to discard. Here, when considering task-positive data, we will use our knowledge of the task timing to sort the ICs into task-relevant and task-irrelevant components in a principled way. To do this, we used a mixed-effect general linear model to estimate which components showed task-dependent temporal variability [24, 60, 139].

The pairing of the ICA and GLM approaches is very powerful for these reasons. The ICA reduces the voxel-level data down to a much smaller component-level description, thus removing the potentially prohibitive number of multiple comparisons. The GLM enables us to identify the relevant ICs in a principled and unbiased way.

6.2.2 Working memory: a very brief introduction

In order to demonstrate the efficacy of our task-positive ICA framework, we selected a cognitive task that is known to recruit multiple, spatially distributed brain regions. Specifically, we implemented a *2-back* working memory experiment. Working memory (WM) describes the mechanism by which information is temporarily cached by the brain for impending use in a cognitive process [6]. Critically, working memory dysfunction has been established as a biomarker for Alzheimer's disease and schizophrenia [6, 48]. There are two key aspects to successful working memory maintenance: *load* and *duration*. If we consider WM as a mechanism for storing discrete items, then *load* refers to the number of items that need to be encoded and retrieved. *Duration* refers to the length of time that the items must be maintained in working memory. The *n-back* task tests WM performance for various loads and durations. Brookes et al. previously investigated WM using MEG in 2011, using a combination of *n-back* and *Sternberg* paradigms to exercise the WM system [28]. They found evidence of theta band (4-8Hz) event-related synchronisation (ERS) due to WM in the frontal regions. They found an associated theta band event-related desynchronisation (ERD) in the visual cortex. Brookes et al. also detected desynchronisations in the high beta/low gamma band (20-40Hz) in the frontal and posterior parietal areas.

These findings demonstrate how WM requires the activation of multiple regions

and across multiple frequency bands. ICA is an ideal tool for investigating the functional connectivity that allows these regions to dynamically integrate in response to the task.

6.3 Materials and methods

6.3.1 Data

Participants: 12 participants took part in this study. The cohort was comprised of six males and six females and had a mean age of 25 years and a standard error of 1 year. The study was approved by the University of Nottingham Medical School Research Ethics Committee.

Paradigm: Each participant was scanned at rest (in a supine position with their eyes open and asked to fixate on a cross) for 5 minutes. Immediately after the resting-state block, the subject underwent a *2-back* working memory task for 12 minutes. The *2-back* paradigm involved the sequential presentation of discrete items (in our paradigm each item was a lower case letter). The subject was asked to respond every time the current item matched the item that was presented two trials before. We refer to this as a **target** trial. Responses were recorded via a button press. The experiment had a block design, consisting of a 30s *2-back* block (made up of 15 trials) followed by a 30s rest/fixation block. Each trial consisted of a 1s blank screen followed by the 1s presentation of a letter. Approximately 3 target trials occurred in each *2-back* block. *PsychoPy* (`www.psychopy.org`) was used to implement the task. Visual presentation was achieved via projection through a waveguide in the magnetically shielded room onto a screen placed approximately 40cm in front of the subject. Figure 6.1 shows a schematic of the *2-back* experiment including the trial-wise structure of each *2-back* block, and the block design of the experiment [88]. The combined block-wise and trial-wise design allowed us to contrast activity between the task and rest blocks and between the target and non-target trials.

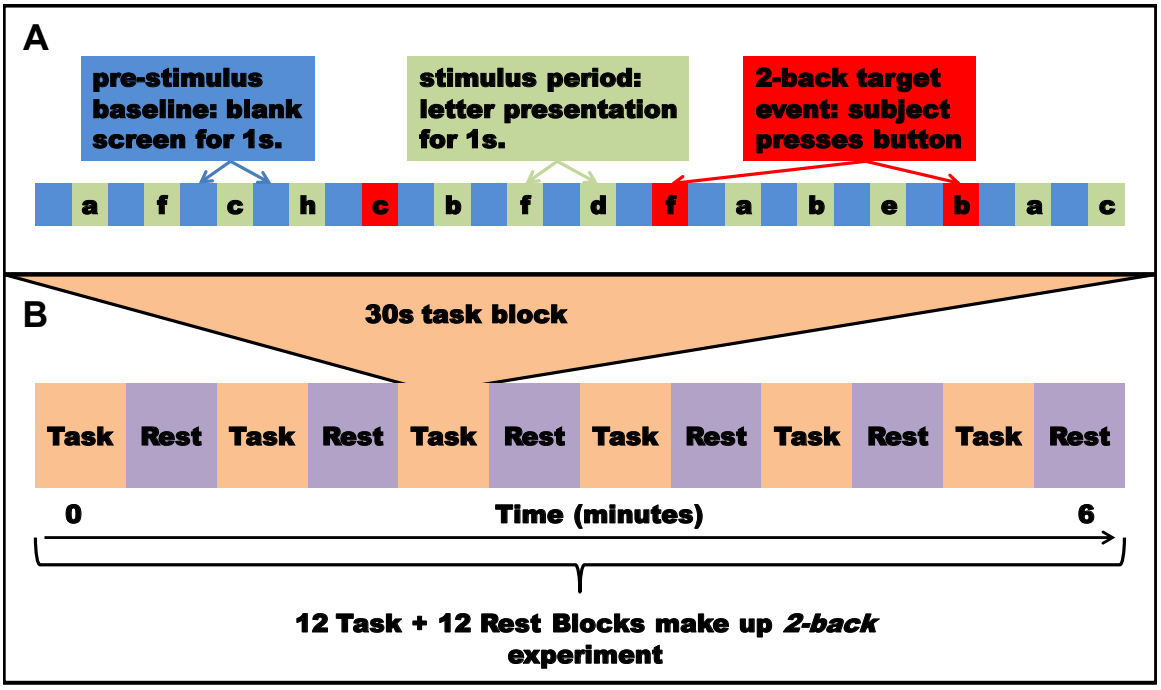


Figure 6.1: *Schematic of the 2-back paradigm. A. The stimulus train: a 1s blank screen preceded a 1s letter presentation to make a single trial. Whenever the current letter matched that of the one two trials earlier, the subject pressed a button. 15 trials made up a 30s task block. B. Each task block was followed by a 30s rest block. 12 task blocks, each followed by a rest block, made up a single experiment, lasting 12 minutes. Figure and caption taken from [88, Fig. 1].*

Data acquisition: The MEG data were acquired using a *CTF* 275 channel whole-head system (*MISL, Coquitlam, Canada*) which sampled the data at 600Hz. We applied 3rd order gradiometer correction to remove far-field interference [136]. We performed head localisation by placing three electromagnetic head-position-indicator (HPI) coils at three locations on the participant's head. The HPI locations and the subject's head shape were measured using a Polhemus Isotrack system. During the scan, the locations of the HPI coils within the scanner were measured by periodically energising the coils. Structural MR images for each subject were acquired using a Philips Achieva 3T MRI system (MPRAGE; 1mm isotropic resolution, 256x256x160 matrix, TR=8.1ms, TE=3.7ms, TI=960ms, shot interval=3s, flip angle=8 degrees and SENSE factor 2).

Data preprocessing: The analysis performed on the *2-back* data was carried out prior to establishing the best-practice preprocessing pipeline (outlined in section 3.6). Instead the data were preprocessed using the following procedure. Channels with

excessively high variance were discarded. Any *2-back* task block/rest block pair that contained abnormally high variance or transient elevations in high frequency power indicative of muscle activity were also discarded.

For each subject, we co-registered their structural MRI to the canonical MNI template using **SPM8** (*FIL, University College, London*). We estimated the lead-fields for each subject using the single shell forward model implemented in **FieldTrip** (*Donders Institute for Brain, Cognition and Behaviour, Radboud University, Nijmegen*) [103]. The data were band-pass filtered into three frequency bands: theta-band - 4-8Hz; alpha/low beta-band - 8-20Hz; high beta/low gamma-band - 20-40Hz. These bands were selected for the following reasons. In section 4.3, we performed an investigation into the optimum *carrier* and *co-variation* frequency window for detecting functional connectivity in the *2-back* data. We found that the 8-20Hz carrier frequency range was the optimum band for measuring functional connectivity (see Figure 4.2B and C). However, previous studies of *n-back* working memory tasks in MEG found that the 4-8Hz and 20-40Hz bands are involved in WM processing [28]. As such, we elected to perform our analysis separately on all three bands.

Source reconstruction for each frequency band was performed using an LCMV beamformer [134, 140]. We regularised the data covariance matrix, $\mathbf{C}$, by boosting its diagonal by four times its smallest eigenvalue (λ_{min}): $\mathbf{C_{regularised}} = \mathbf{C} + 4 \times \lambda_{min} \times \mathbf{I}$, where $\mathbf{I}$ is an $N_{sensors} \times N_{sensors}$ identity matrix. Source space estimates of neural activity were estimated for every vertex of a 6mm grid spanning the whole brain. In addition, eye-blink time courses were estimated by beamforming at each eyeball (MNI coordinates: $[\pm 34, 56, -38]$mm).

6.3.2 Imaging functionally-connected networks

Each subject's band-limited, source-space estimate of neural activity at every voxel was *weights-normalised* (which we demonstrated was an essential preprocessing step in section 4.5.2.2). For each voxel, the oscillatory amplitude envelope was estimated by taking the absolute of the *analytic signal*, computed via the Hilbert transform. These envelope time courses were low-pass filtered using a 0.5s windowed-average

down-sampling. Although we previously demonstrated that the optimal window length (Δ) is in the range 1s$<\Delta<$4s, such aggressive down-sampling would render our analysis insensitive to within-trial fluctuations in band-limited power (as the trials all have a 2s length). By reducing Δ to 0.5s, we estimated 4 samples per trial and retained some sensitivity to within-trial activity.

We performed spatial smoothing (using a 4mm full-width-at-half-maximum Gaussian kernel), spatial re-sampling (to an 8mm grid), de-meaning and temporal concatenation in exactly the same way as we did previously for our resting-state studies in chapters 4 and 5. However, we included the beamformed signals of the activity originating from the eye-balls in the concatenation in order to encourage the ICA to find a single eye-blink related component.

When performing the ICA decomposition, we used the **ICASSO** algorithm[1] instead of the standard **FastICA** package [73]. **ICASSO** is an extension of the **FastICA** software that attempts to overcome the issue of local maxima. ICA algorithms start with a random guess of the mixing matrix and then perform an optimisation from there to maximise some measure of *non-Gaussianity* across the components. However, they can find sub-optimal decompositions if the optimisation gets stuck in a local maximum. **ICASSO** runs multiple **FastICA** decompositions (e.g. typically about 30 runs) and uses a clustering algorithm to group the matching components across the different runs. The tightness of the resulting clusters is an indicator of component consistency. Furthermore, the cluster centre can be used as an alternative estimate of the optimum set of independent components. One potential downside of **ICASSO** is that the cluster centre may represent a less optimal ICA decomposition than some of the single **FastICA** runs used to estimate it (particularly if there are many local maxima corresponding to substantially different levels of non-Gaussianity). Furthermore, using **ICASSO** comes with an additional computational cost as it has to perform multiple runs of *FastICA*.

We performed 6 ICAs using **ICASSO**. For each frequency band, we input the concatenated *combined task and rest* blocks for one ICA and just the *task blocks only* for the second ICA (we will use *combined task and rest* and *task blocks only* to denote

thee two types of analysis in the remainder of this chapter). We performed these two distinct analyses to demonstrate that our task-positive ICA framework can be applied to both block design and trial design experimental paradigms. For each ICA run, we computed 20 independent components and estimated the map of Pearson correlation coefficients between the ICs and the concatenated envelope time courses input to the ICA at each voxel.

6.3.3 Statistical framework for identifying *2-back* relevant FCNs

Each of the 6 ICAs produced 20 independent time courses from which we estimated 20 corresponding correlation maps (refer to section 4.5.2.4 for a description of correlation maps as a way to visualise RSNs). We used a mixed-effects GLM to test each *combined task and rest* ICA for components that showed strong block-dependent time courses. We used an equivalent mixed-effects GLM on the *task blocks only* ICAs to detect components which showed significant differences between *2-back* target trials and non-target trials. We contrasted these component-level analyses with voxel-level mixed-effects GLMs in order to demonstrate the improvement that our combined ICA-GLM framework provides. Specifically, for each frequency band we performed four statistical analyses:

1. Voxel-wise mixed effects GLM on the *combined task and rest* blocks.
2. Voxel-wise mixed effects GLM on the *task blocks only*.
3. Component-wise mixed effects GLM on the *combined task and rest* blocks.
4. Component-wise mixed effects GLM on the *task blocks only*.

The mixed-effects GLM: The general linear model is a framework for quantifying the linear dependencies between the external, task-related stimuli of a given paradigm and the recorded neural activity. The mixed-effects GLM is designed to model the variability of response of neural activity to the task at the subject-level (i.e. how

[1]An open-source Matlab implementation of **ICASSO** can be downloaded from `http://research.ics.aalto.fi/ica/icasso/`.

much variability is there in the strength of response for a given voxel for one subject across multiple trials) and at the group-level (i.e. how much variability is there in the strength of response to a stimulus across multiple subjects). The mixed-effects GLM accounts for both sources of variability by performing separate subject-level GLMs (each modelling the variability specific to that subject), followed by an overall group-level GLM (that models the variability across subjects).

The mixed-effects GLM for contrasting task and rest blocks: We produced two regressors (box-car, column vector time series of equal number of samples to our windowed-averaged envelopes) for each subject's data. The task block regressor had a value of 1 for samples that occur within the task blocks and 0 otherwise. Conversely, the rest block regressor had a value of 1 during the rest blocks and 0 elsewhere. We also included the mean of the windowed-average down-sampled envelopes of the source-reconstructed eye-ball time courses as a confound regressor. We modelled the windowed-averaged envelope at each voxel as a weighted sum of these three regressors. Similarly, we modelled the subject-specific segment of each independent time course from the *combined task and rest* ICA using the same GLM. We show a schematic of the *combined task and rest* GLM in Figure 6.2A. Mathematically, we express our subject-level GLM as:

$$\mathbf{Y_n} = [\mathbf{X_{Task}}\mathbf{X_{Rest}}\mathbf{X_{Eyes}}]\mathbf{B} + \boldsymbol{\varepsilon} \tag{6.1}$$

where $\mathbf{X_{Task}}$, $\mathbf{X_{Rest}}$, and $\mathbf{X_{Eyes}}$ are the task block, rest block and eye-blink confound regressors respectively (shown in Figure 6.2A). $\mathbf{Y_n}$ is the down-sampled envelope for the n^{th} subject at a single voxel (for a voxel-wise analysis) or at a single component (for a component-wise analysis). $\mathbf{B}$ is the 3-element vector of weights or parameter estimates ($\mathbf{B} = [\beta_{Task}, \beta_{Rest}, \beta_{Eyes}]^T$). As we were interested in voxels/components that show differences between task and rest blocks, we estimated the contrast between parameter estimates (COPE) for the task and rest blocks ($\beta_{Task-Rest} = \beta_{Task} - \beta_{Rest}$). As we included the eye-blink activity as a confound regressor, our COPEs are

orthogonal to the eye-blink activity and so any statistical findings cannot be explained as an eye-blink artefact.

We performed subject-level GLMs across every component and across every voxel. We then performed a group-level GLM on the COPES from each subject at every voxel/component. We modelled the vector of 12 COPEs, $\mathbf{B}_{Task-Rest}$, (one for each subject) as the product of a 12-element column vector of ones, $\mathbf{Z}$, and the group-level parameter estimate, β plus a group-level normally distributed error $\boldsymbol{\mu}$:

$$\mathbf{B}_{Task-Rest} = \mathbf{Z}\beta + \boldsymbol{\mu} \tag{6.2}$$

The group-level stage provided a statistical parametric map showing statistical differences between task blocks and rest blocks for the voxel-level GLM. For the component-level GLM, we estimated a vector of statistics, each element corresponding to an independent component time course/correlation map pair.

The mixed-effects GLM for contrasting target and non-target trials: We produced two regressors (box-car, column vector time series of equal number of samples to our windowed-averaged envelopes for *the task-blocks only* envelopes) for each subject's data. The target-trial regressor had a value of 1 for samples that occur within the *2-back* target trials and 0 otherwise. Conversely, the non-target regressor had a value of 1 during the non-target trials and 0 elsewhere. Again, we included the eye-blink envelope time course as a confound regressor. We modelled the windowed-averaged envelope at each voxel/component as a weighted sum of these three regressors. We performed equivalent subject level GLMs for every voxel and every component in the *task blocks only* data (shown in Figure 6.2B):

$$\mathbf{Y_n} = [\mathbf{X_{Target}}\mathbf{X_{Non-target}}\mathbf{X_{Eyes}}]\mathbf{B} + \boldsymbol{\varepsilon} \tag{6.3}$$

where $\mathbf{X_{Target}}$, $\mathbf{X_{Non-target}}$, and $\mathbf{X_{Eyes}}$ are the target trial, non-target trial block and

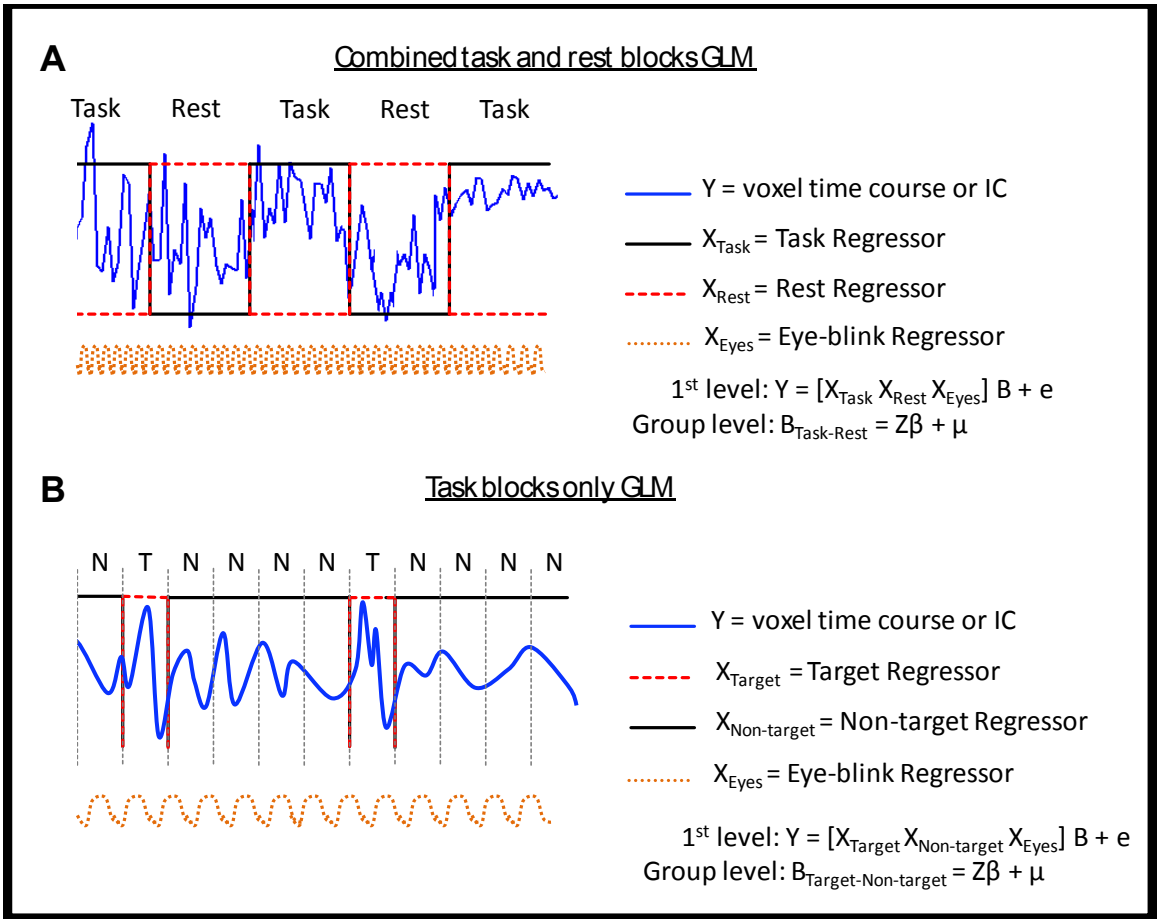


Figure 6.2: *Testing of independent components estimated from the 2-back data using a mixed-effects GLM. A. The combined task and rest mixed-effects GLM, testing for differences between the task blocks and rest blocks. B. The task blocks only mixed-effects GLM, testing for differences between the target (T) and non-target (N) trials. Both GLMs include the average eye-blink activity for each subject as a confound regressor in the 1st level, ensuring that any statistical inferences are orthogonal to eye-blink activity. Figure and caption taken from [88, Fig. 2].*

eye-blink confound regressors respectively and where $\mathbf{B} = [\beta_{Target}, \beta_{Non-target}, \beta_{Eyes}]^T$. We then performed an identical group-level GLM on the vector of COPEs contrasting target and non-target trials, $\mathbf{B}_{Target-Non-target}$:

$$\mathbf{B}_{Target-Non-target} = \mathbf{Z}\beta + \boldsymbol{\mu} \tag{6.4}$$

to give a statistical map across all voxels showing differences in envelope activity during target trials compared with non-target trials.

6.3.4 Correcting for multiple comparisons

We used the false discovery rate method (FDR) to correct for multiple comparisons across voxels in the statistical maps for the *combined task and rest* and *task blocks only* voxel-wise GLM analyses [19]. FDR determines the appropriate statistical threshold that corresponds to a specific expected proportion of false positives (e.g. 5%) [101]. FDR corrections were implemented using **FSL** (*FMRIB,Oxford*).

In order to correct for multiple comparisons across components in our component-wise analyses, we utilised a multi-step-up test [101]. For a given ICA, we ranked the ICs in order of significance (from least to most significant). Starting with the least significant component, we tested it against the Bonferonni-corrected threshold (i.e. when testing 20 components with a significance threshold $\alpha = 0.05$, we tested if $p \leq \frac{0.05}{20}$). If it failed to pass significance, the next least significant component was tested against the same threshold (step 2), scaled up by the step number $p \leq \frac{2 \times 0.05}{20}$. If it proved to be non-significant, the next component was tested until one was found to be significant. All remaining (more significant) components were then classed as above threshold. We present the two rules of the multi-step-up test below:

if $p_i > \frac{\alpha i}{I}$, **then** test next component (increment i).

if $p_i \leq \frac{\alpha i}{I}$, **then** classify components i through I as significant.

We used the multi-step-up test as it is equivalent to the FDR correction that we used in the voxel-wise analyses [101].

6.4 Results

Voxel-wise GLMs: Figure 6.3 shows the Z-statistical maps from the 6 voxel-wise GLM analyses (for A. the *combined task and rest* envelopes and for B. the *task blocks only* envelopes, over three frequency bands each: 4-8Hz, 8-20Hz, 20-40Hz). Each map shows corrected Z-statistics. For Figure 6.3A all Z-statistical maps were thresholded between 3 and 5. In the 4-8Hz band, we found significant activity in the frontal, visual and motor regions. In the 8-20Hz band, we found significant activity in the parietal cortices and superior parietal lobule but this activation was anatomically unspecific.

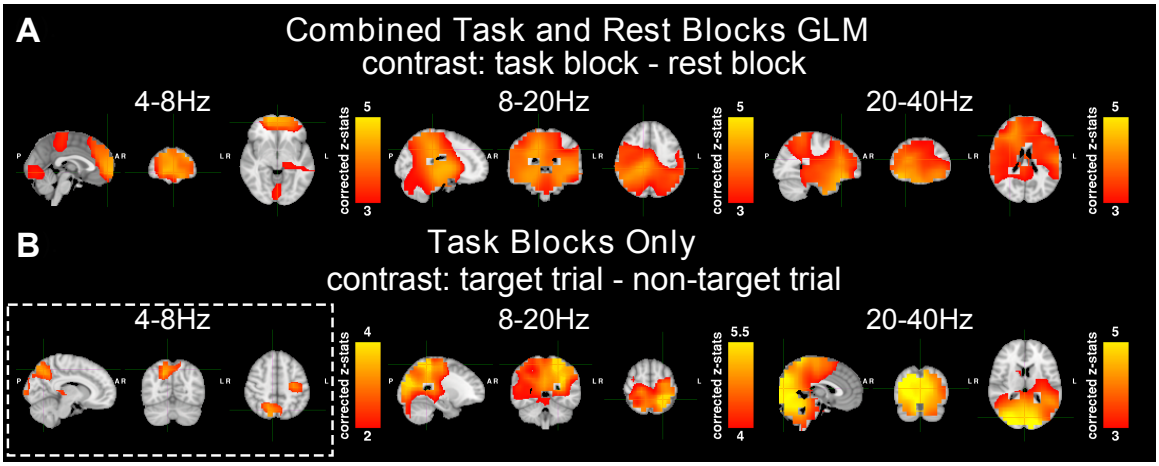


Figure 6.3: *Results of mixed-effects GLM analyses on the down-sampled oscillatory envelopes at each voxel. Maps of corrected Z-statistics are shown in radiological view (axial, coronal and sagittal views from left to right). Z-statistics were corrected using false discovery rate (FDR) methods in* ***FSL*** *and subsequent Z-statistical maps thresholded between 3 and 5 for visualisation, except the task blocks only 4-8Hz (2<Z<4) and 8-20Hz (4<Z<5.5) maps. An FDR of 0.05 was used for all the GLMs except the 4-8Hz task blocks only analysis (in the white dashed box) where a FDR of 0.2 was used as no activity survived corrections at a FDR of 0.05. A. Result of the combined task and rest GLMs testing for differences between the 30s task and rest blocks for the 4-8Hz, 8-20Hz and 20-40Hz bands. B. Result of the task blocks only GLM testing for changes between the trials where a target was presented against all other non-target trials. In all GLMs, the average eye-blink time course was included as a confound regressor and hence all statistical inferences are orthogonal to eye-blink activity. Figure and caption taken from [88, Fig. 4].*

In the 20-40Hz band, we found activity in the parietal, frontal and temporal regions.

In Figure 6.3B we present the Z-statistical maps from the *task blocks only* voxel-wise analyses. We modified the thresholds in the maps to improve visualisation (for 4-8Hz: 2<Z<4; for 8-20Hz: 4<Z<5.5; for 20-40Hz: 3<Z<5). In order to detect any significant activity in the 4-8Hz band, we had to raise the FDR (the expected rate of false positives) from 0.05 to 0.2. When contrasting target trials with non-target trials, we found significant activity in the visual cortex, medial superior parietal lobule and the left motor cortex. In the 8-20Hz band, we found significant activity in the parietal regions. In the 20-40Hz band, we found significant differences in the parietal lobes, visual areas and cerebellum.

ICA/GLM results - *combined task and rest* analyses: Figure 6.4 shows 5 correlation maps (out of 20) for each of the three ICAs performed on the envelopes of band-limited neural activity from the *combined task and rest* blocks concatenated

over subjects (along with their corresponding significance value). We found more than five significant ($p_{corrected} \leq 0.05$) components for each band but have only displayed the most significant/relevant for clarity.

Theta band (4-8Hz): In the theta band, we found components that corresponded to two frontal networks, a visual network, the left hippocampus and the SPL. The two frontal components agree with our voxel-wise analysis but the ICA has split the task-related activity into two temporally distinct components. The visual network and SPL also agree with our voxel-wise GLM.

Alpha/low beta band (8-20Hz): We identified five components corresponding to the left and right superior temporal gyrus, bilateral SPL, DMN and bilateral insula.

High beta/low gamma band (20-40Hz): We found components corresponding to the DMN, somatosensory network, anterior cingulate cortex, SPL and bilateral hippocampi.

ICA/GLM results - *task blocks only* analyses: Figure 6.5 shows the results of the three ICAs performed on the oscillatory envelopes of the band limited neural activity of the *task blocks only*, temporally concatenated across subjects. Again, we selected 5 correlation maps from each ICA. Unlike the *combined task and rest* analyses, fewer than five components per frequency band showed significant target versus non-target changes in activity. As such, we also selected non-significant components that had functionally interesting spatial maps.

Theta band (4-8Hz): We found three significant ($p_{corrected} \leq 0.05$) components corresponding to the right motor cortex and two frontal networks (which were extremely similar to the two significant frontal components found in the theta band *combined task and rest* blocks ICA in Figure 6.4). Two non-significant components were found corresponding to the SPL and lateral visual networks.

Alpha/low beta band (8-20Hz): Two significant components were found corresponding to the bilateral insula and the cerebellum. We also found three non-significant components that corresponded to the left frontoparietal network, the DMN and the left motor cortex.

High beta/low gamma band (20-40Hz): We found two significant components for the visual network and bilateral hippocampi. We also found a near-significant ($p_{corrected} = 0.055$) component corresponding to the pre-motor network. Again we found non-significant left motor cortex and SPL components.

6.5 Discussion

In this chapter, we presented an adaptation of our ICA-based analysis for extracting resting-state networks. The adapted analysis uses ICA to extract functionally-connected networks from task-positive MEG data. We combined the output of the ICA with a mixed-effects GLM to sort through the independent components and identify any FCNs that demonstrate task-dependent activity. We demonstrated this analysis on a *2-back* working memory experiment and contrasted it with a more typical voxel-wise mixed-effects GLM (also performed on the windowed-averaged envelopes). Furthermore, we used the framework of the GLM to include eye-blink activity (estimated by beamforming at both eye-balls) as a confound regressor. By doing this, we ensured that all our statistical inferences were orthogonal to the estimated eye-blink activity. Any confound regressor could be included in this framework, allowing us to account for any non-neural artefact that might affect our analysis, so long as we can estimate a temporal regressor for the artefact.

Task-positive ICs and working memory: From both the task-positive ICAs performed on the *combined task and rest blocks* and the *task blocks only*, we found a selection of significantly task-correlated FCNs corresponding to regions which are known to be associated with working memory processing. Specifically, we found two theta-band frontal components. Changes in frontal theta oscillations have consistently been associated with working memory [28, 83, 107] as well as with other cognitive functions [80]. We also found frontal theta changes in the voxel-wise GLM analysis contrasting task and rest blocks. However, we found no significant theta band differences in the voxel-wise analysis contrasting target and non-target trials.

That we found significant frontal theta at the component-level but not the voxel-level demonstrates one of the benefits of the ICA in improving the sensitivity of the subsequent statistics.

It is noteworthy that we found two frontal theta components rather than a single one. One possibility is that this reflects two functionally distinct processes occurring in response to the *2-back* experiment. However, this could be an artefact of the ICA model order selection (refer to section 4.5.1.4). In order to establish this, we would need to run multiple ICAs at lower and higher model orders to see if the two components merge or split up [1]. Performing such an analysis was beyond the goal of this study, but would be important in any future investigation into frontal-theta components.

We also found significant theta-band differences in the visual cortex, high beta/low gamma band differences in the parietal regions, alpha/low beta band differences in the left and right superior temporal gyri, probably corresponding to the temporoparietal junction network (TPJ): all of these regions have been independently shown to respond to working memory function in those frequency bands [26, 28].

We have validated our task-positive ICA extensively by accurately localising working memory related activity both in space and in frequency. This demonstrates that the task-positive ICA framework can be used to map the rich spatial/spectral characteristics of complex cognitive tasks.

Imaging functionally-connected networks: In both the *combined task and rest* and *task blocks only* ICAs we found several FCNs that correspond to RSNs that have previously been imaged in MEG and fMRI (refer to section 4.5 for examples of such RSNs). These FCNs include the visual network, the default mode network, the bilateral insula (which we postulate might be part of the salience network), the sensorimotor network and the somatosensory network.

Most interesting are the DMN and bilateral insula. The DMN has been shown to be a task-negative network, "switching off" during goal-directed activities [56, 108]. It is very interesting that we find the DMN time course to be significantly correlated

with the task-rest block time course. We extracted the DMN from the *task blocks only* ICA as well and found that it showed no significant changes between target and non-target trial, which is plausible if it is in a down-regulated, "off" state during the whole task block.

If we are correct in asserting that the bilateral insula component is a sub-network of the salience network, then the simultaneous extraction of both these networks from a single ICA is of great interest as the salience network has been suggested to be involved in switching between the DMN/task-negative state and other brain states [125].

Block-wise design versus trial-wise design: We demonstrated that our task-positive ICA framework can be applied to both block-wise experimental designs (common in fMRI) and trial-wise designs (more common in EEG/MEG studies). Our *2-back* paradigm had both a block-wise design and trial-wise design (if we only considered the *2-back* blocks). As such, we could treat the data in both ways, which we did by performing both *combined task and rest* analyses and *task blocks only* analyses. As we would expect, we found many similar components in both approaches (unsurprising as the *task blocks only* data is a 50% subset of the *combined task and rest* data).

The major difference between the two approaches became apparent when performing statistical inferences on the resulting components. The contrast in neural activity during rest and during the *2-back* task is extremely strong, as evidenced by the fact that we found more that 5 significant ICs per ICA. In contrast, we only found 7 significant components across the three *task blocks only* ICAs. This is most likely a reflection of the much subtler differences between the target trials and non-target trials. In this experimental design, the trial-wise approach enables us to ask many more questions beyond target-vs-non-target. For instance, we could look at correctly-identified *2-back* targets versus incorrectly-identified, even focusing on the pre-stimulus period to see if we can predict correct trials (assuming we recorded enough data). Such exotic analyses are beyond the scope of research of this thesis

and we simply wish to acknowledge the plethora of interesting questions that can be addressed with the trial-wise analysis.[2] We note that the reduced sensitivity of the trial-wise analysis is in part possibly due to reduced amount of data. However, we think this is unlikely since the number of samples in the *task blocks only* analysis was well over that used in our resting-state studies.

When performing task-positive ICA, our windowed-average down-sampling of the oscillatory envelope presents us with the following trade-off: as we showed in Figure 4.2, greater down-sampling improves our ability to detect functional connectivity. However, greater down-sampling comes at the cost of within-trial resolution. When using a window length of 0.5s to investigate changes in FCNs between task and rest blocks, we are not exposed to any such issues as each block is well-sampled with 60 data points. However, when considering the trial-wise analyses, each trial is only sampled with 4 data points and this may explain the weaker statistics associated with the *task blocks only* networks.

There is one potential pitfall with our task-positive ICA framework - that is our assumption that the sources we wish to recover are temporally independent. In the resting-state, temporal independence between the underlying network processing is reasonable. As we use external task stimuli to lock the brain's activity to a single process, we may remove the independence that we use in the resting-state to separate the processes associated with different networks. In our *combined task and rest* analysis, the spontaneous activity of the rest blocks protects us from violating our independence assumptions. In the *task blocks only* analysis, although we do not have equivalent protection, the correspondence between the MEG RSNs found in chapter 4, the *combined task and rest* FCNs and *the task blocks only* FCNs suggests that our independence assumption is not being violated.

Detecting task-correlated activity in the hippocampi: We identified components that spatially corresponded to activity in the area of hippocampus. Specifically,

[2]We have utilised the trial-wise task positive ICA/GLM framework to investigate the role of the right frontoparietal network in determining successful visual short term memory encoding [4, Under review].

we found theta band hippocampal components in the *combined task and rest* ICA and high beta/low gamma components in both the *combined task and rest* and *task blocks only* ICAs. All these components were found to be significantly task-related ($p_{corrected} < 0.05$).

This result is noteworthy for two reasons. Firstly, it further validates our analysis as theta and gamma band activity have previously been implicated in working memory function [5, 128]. Secondly, it provides evidence that hippocampal sources can be imaged with MEG. In chapter 2, we raised the question of whether activity from sub-cortical sources, such as the hippocampus, is detectable with MEG. This is because sub-cortical sources typically do not have the columnar cyto-architecture which generates open fields and contributes to a measurable MEG signal. Here, we have found strong evidence that we can in fact detect hippocampal activity. Quraan et al. demonstrated with simulations that hippocampal sources could be detectable under the following conditions: 1.) a suitable paradigm is used to stimulate the hippocampi; 2.) sufficient numbers of trials are recorded (150 trials for 12 subjects); 3.) a suitable source reconstruction and analysis approach is used. As we have previously stated, several studies (Tesche and Karhu in 2000 and Axmacher et al. in 2011) have previously demonstrated hippocampal involvement in WM [5, 128]. Our task data meets the necessary criteria of both sufficient numbers of trials and subjects. Furthermore, our analysis framework is well-suited for detecting hippocampal activity. At the source reconstruction stages, we band-pass filter into the theta and high/beta low gamma bands, removing a large amount of interference and noise. Our beamformer has a high degree of spatial specificity and so provides an additional layer of interference suppression. In source space, the ICA is ideally suited to separate the weak band-limited envelope signal associated with the hippocampi even in the presence of stronger sources of interference and noise (this assumes that our PCA dimensionality reduction has not removed the weak signal). Finally, the GLM demonstrates that the hippocampal components have task-locked activity and are not just noise components with spatial maps that coincidentally centre on the hippocampi. We also used the GLM to discount signal leakage from the eye-blink artefact by in-

cluding it as a confound regressor. We do concede that our hippocampal correlation maps lack the spatial resolution to confirm that the neuronal sources underpinning the hippocampal ICs definitely originate from the hippocampi and not neighbouring temporal lobe cortex.

Advantaged of the ICA/GLM combination: When using ICA, the components are generated in an arbitrary order. This happens even under idealised circumstances (i.e. in the absence of noise and when the true model order is known). The consequence is that components of interest are mixed together with irrelevant components. In our resting-state ICAs (in section 4.5) we identify components of interest by matching them to fMRI-derived RSNs. This is one example of the sorts of heuristic approaches which are often used to decide which components are of interest and which are irrelevant. Such approaches are consequently biased to classifying only what is expected *a priori* as of interest. In our framework, by using the GLM to infer which components demonstrate a task dependency and which do not, we have developed a principled and unbiased solution to this issue. Of course, this is only possible in task-positive data where we can define temporal regressors that correspond to the task timing.

Conversely, mass-univariate statistical approaches such as the GLM are limited by the issue of multiple comparisons. In order to perform a fully exploratory analysis at every voxel, we must account for the thousands of statistical tests performed [101,139]. Even accounting for dependence between voxels, whole-brain analyses can be prohibitively large for small effects [11]. One solution is the *a priori* selection of "regions of interest" (ROIs). Whilst ROI selection ameliorates the multiple comparisons issue, it again exposes us to only finding that which we expect to find. By using the ICA to reduce the data, we can still perform whole-brain investigations without being overwhelmed by multiple comparisons. As ICA is a data-driven approach, we are not imposing our prior expectations but instead letting the features of the data determine the reduced representation (with the caveat that by using ICA, we are imposing the constraint of statistical independence which may not be the best feature with which to decompose our data).

6.6 Summary

In this chapter, we presented a method for using ICA to extract functionally-connected networks from task-positive MEG data and for using a mixed-effects GLM to perform statistical inference on the resulting ICs. We demonstrated this approach on a *2-back* working memory paradigm and validated it against a voxel-wise GLM approach. We showed that the task-positive ICA combined with the GLM was able to locate task-correlated activity with excellent correspondence (in both space and frequency) to the voxel-wise GLM approach and to other working memory studies. Furthermore, we demonstrated that our approach is flexible and can be used to analyse experiments that have block-wise designs or trial-wise designs. We demonstrated our approach as an exploratory and unbiased method for investigating functional connectivity in task-positive data which accounts for multiple comparisons in an efficient but principled manner. Finally, we demonstrated that pairing ICA with the GLM provides an approach that can potentially outperform either method in isolation.

In the previous chapter, when considering MEG-adapted dual regression (MADR), we focused exclusively on detecting differences in resting-state networks. In this chapter, we focused on how to best utilise our knowledge of the task's design, specifically its temporal structure, in the context of ICA-based functional connectivity analysis. That said, the MADR analyses could be applied to task-positive data in an identical fashion. Given a data set consisting of task-positive MEG data from multiple sub-groups, we could employ MADR to extract group-level component time courses (using any spatial basis set). We could then utilise our task-positive analysis to classify the task-correlated and task-irrelevant components. Finally, we could use the MADR to look for differences in the sub-group maps and time courses. Our *2-back* cohort is only made up of a single group and so it is not possible to do such an analysis in this instance.

One limitation of our analysis as currently framed is that we have sacrificed the fantastic temporal resolution afforded to us by MEG in order to extract functional networks. However, it is possible to return to the millisecond temporal resolution

of raw MEG data. Brookes et al. (2012) did this by estimating FCNs on the 2s windowed-averaged data via a temporal ICA and then using the FCN maps as ROIs to perform time-frequency decompositions [26]. Such approaches will be essential if we are to capitalise on all the information that task-positive MEG has to offer.

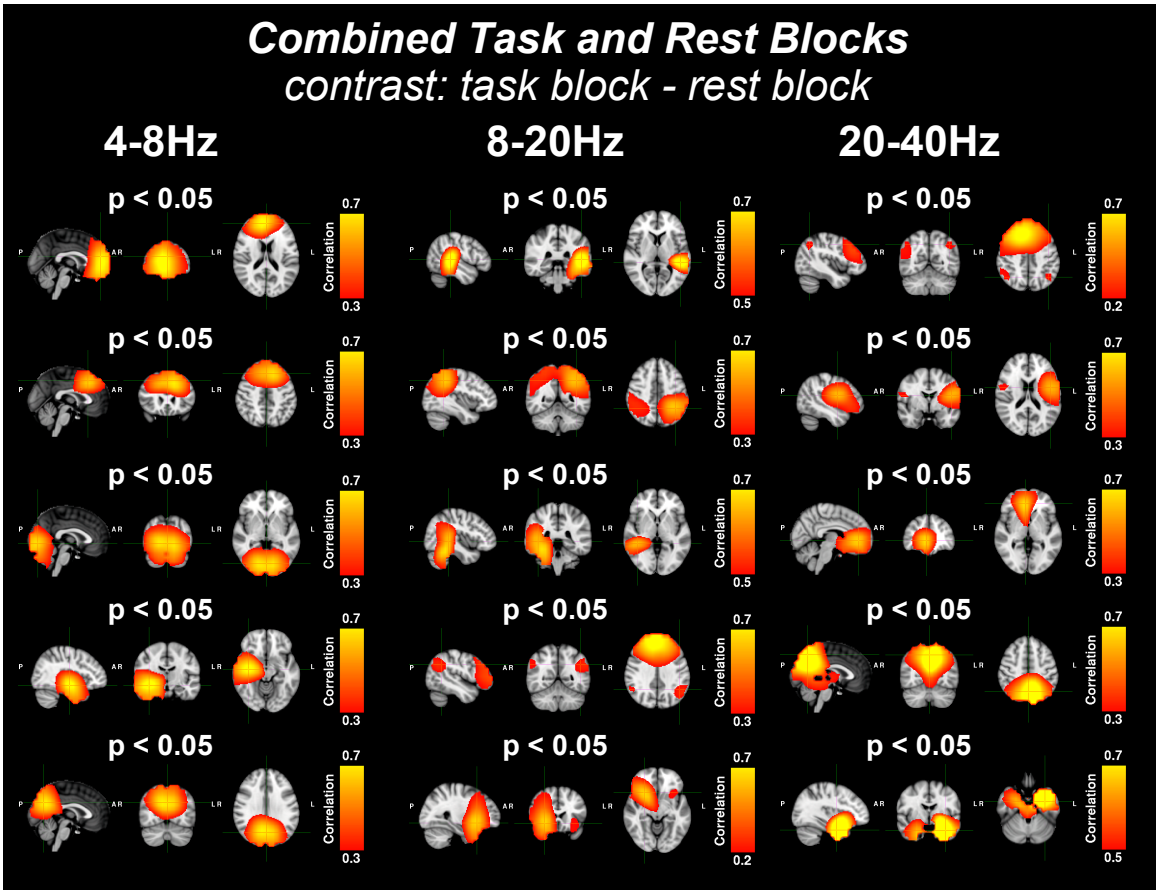


Figure 6.4: *This figure shows the correlation maps generated from ICAs applied to the combined task and rest blocks for the down-sampled envelopes of three frequency bands (4-8Hz, 8-20Hz, 20-40Hz), concatenated over subjects. Correlation maps, presented in radiological view (axial, coronal and sagittal slices from left to right), show the correlation between each temporally independent component and the concatenated down-sampled envelopes of each voxel. Variable correlation thresholds were used to aid visualisation and are shown on each map's colour bar. For each frequency band, five correlation maps are presented for components whose corrected p-values were found to be significant ($p_{corrected} \leq 0.05$). From top to bottom, maps correspond to: 4-8Hz: frontal area; separate frontal area, visual cortex, right hippocampus, superior parietal lobule; 8-20Hz: left superior temporal gyrus, bilateral superior parietal lobule, right superior temporal gyrus, default mode network, bilateral insula; 20-40Hz: default mode network, somatosensory network, cingulate, medial superior parietal lobule, bilateral hippocampi. Figure and caption taken from [88, Fig. 5].*

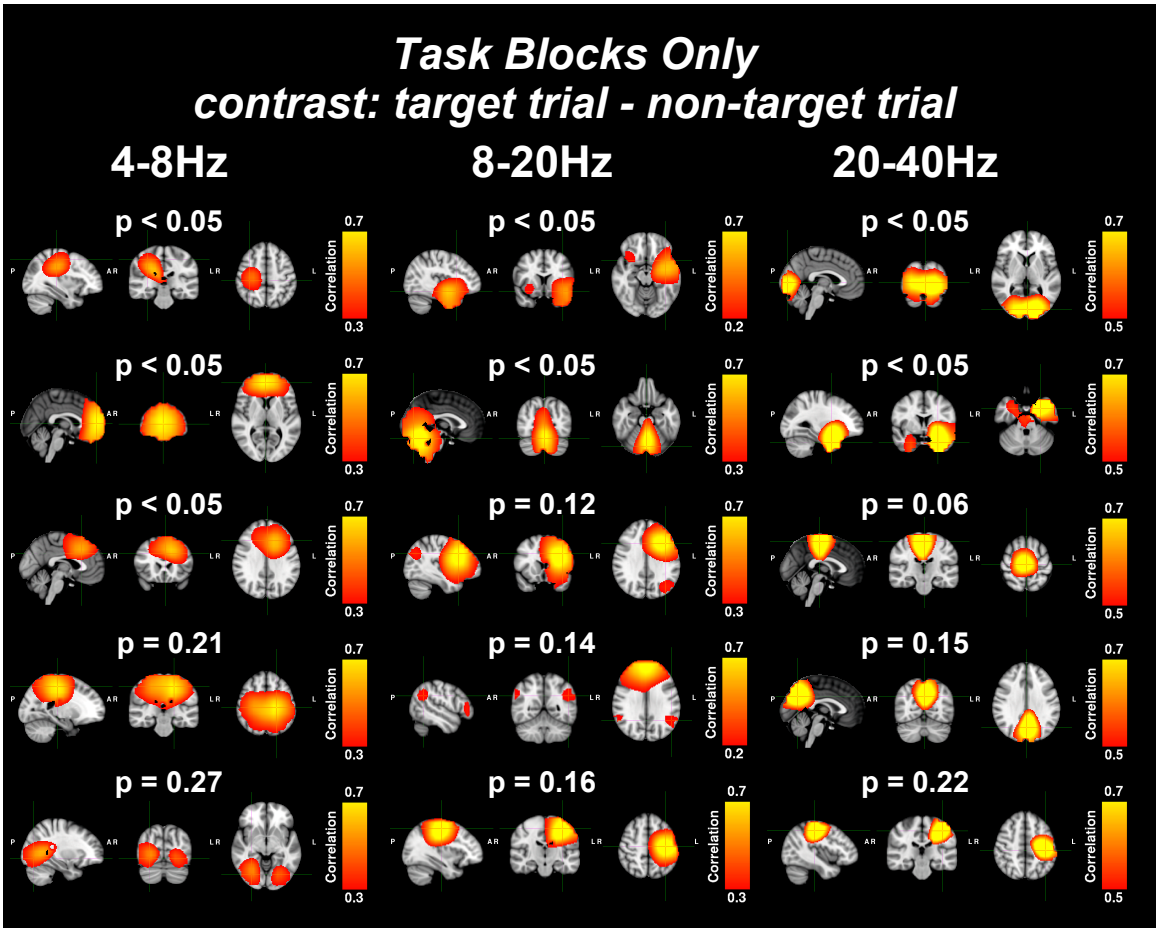


Figure 6.5: *This figure shows the correlation maps generated from ICAs applied to the task blocks only for the down-sampled envelopes of three frequency bands (4-8Hz, 8-20Hz, 20-40Hz), concatenated over subjects. Correlation maps, presented in radiological view (axial, coronal and sagittal slices from left to right), show the correlation between each temporally independent component and the concatenated down-sampled envelopes of each voxel. Variable correlation thresholds were used to aid visualisation and are shown on each map's colour bar. For each frequency band, five correlation maps are presented for components whose corrected p-values are significant ($p < 0.05$) or whose spatial localisation is functionally interesting. From top to bottom, maps correspond to: 4-8Hz: right motor cortex, frontal areas, separate frontal areas, parietal cortex, lateral visual areas; 8-20Hz: bilateral insula, medial cerebellum, left lateral frontoparietal network, default mode network, left motor cortex; 20-40Hz: visual cortex, bilateral hippocampi, premotor areas, medial superior parietal lobule, left motor cortex. Figure and caption taken from [88, Fig. 6].*

Chapter 7

Parcellation and network analyses for MEG

7.1 Chapter abstract

In this chapter, we present a framework for estimating a data-driven parcellation for MEG, using spatial ICA, and for performing a partial correlation network analysis to characterise the functional network structure between parcels. In particular, we focus on developing a novel method for removing the signal leakage between a set of parcels that is compatible with multi-parcel network modelling methods such as partial correlation. We validate our analysis by estimating the standard and partial cross-frequency correlation matrices for the slow envelope oscillations of the band-limited neural activity of each parcel. We present preliminary evidence for the existence of stationary, direct FC between the slow envelope oscillations of different frequency bands, both within and between parcels.

7.2 Introduction

We have so far considered resting-state (and task-positive) networks as individual, homogeneous functional units of brain activity. While these kinds of representations are extremely useful for grouping brain regions into functionally-related clusters, they

typically do not yield much information about the internal structure of functionally-connected networks. Network analysis techniques have emerged as an approach that can reveal this structure.

7.2.1 Network analysis

Network analysis can be broken down into three key stages: defining a parcellation, estimating the network matrix, and analysing the network matrix.

1. Defining a parcellation: A parcellation is a segmentation of the whole brain into a number of individual regions, such that it is reasonable to model all the neural activity originating within each parcel as a single time course. Ideally, the distribution of parcel volumes and the number of parcels will be inversely related (i.e. doubling the number of parcels roughly halves the size of the resulting parcels rather than preserving some large parcels but producing additional tiny ones). Furthermore, the distribution of parcel sizes should be restricted to a sensible range (i.e. we do not want one parcel that corresponds to an entire lobe whilst another corresponds to just a few voxels). Finally, an ideal parcellation should be made up of bilateral pairs of contiguous parcels in order to mirror the observed symmetry of RSNs. The parcellation should also respect boundaries in the underlying anatomy of the brain (such as the transition between grey and white matter).

Here we outline three possible techniques for parcellating the brain. However, this is very much an active research area and these techniques do not represent the final word in parcellation methods. The first approach to parcellating the brain is anatomically motivated. By delineating regions based on changes in anatomical features, one can separate regions with differing anatomy [131]. Anatomical parcellations are motivated by the idea that anatomical structure must drive function [45] and so anatomically homogeneous regions should have some functional homogeneity. However, in the context of fMRI, Smith et al. (2011) showed, by considering network analyses on simulated fMRI data, that anatomical parcellations tend to perform poorly because they blur functionally distinct network nodes together. This distorts the resulting

inferred network structure. As such, Smith et al. cautioned against their use [124].

The second approach we consider is known as edge-based parcellation [36]. By considering the correlation matrix between all voxels, regions can be extracted within which the resting-state functional connectivity is homogeneous. Furthermore, sharp transitions in functional connectivity can be extracted and used as parcel boundaries. The advantage of this technique is that it is by definition constructing edges between parcels based on changes in the homogeneity of functional connectivity. However, it does not directly output a usable parcellation and requires further stages of image processing which introduce complexity.

The final technique we consider is to use a high-dimensional group spatial ICA to estimate the functional parcellation. As described in section 4.5.2.3, spatial ICA estimates a set of spatially non-overlapping 3D maps and associated time courses. In fMRI, when the ICA model order is sufficiently low (approximately ≤ 30), these maps correspond to whole networks. As the model order increases, the networks get split up, first into sub-networks and then into single nodes [1]. In fMRI, ICA model orders of approximately 100 are used to define a group average parcellation. Spatial ICA has the following advantages as a parcellation technique. The user has very clear control (via the model order) of the scale of the parcellation. Secondly, the high-dimensional parcellation can be related to RSNs from a low-dimensional ICA on the same data. Finally, group ICA can be used to enforce consistency in the parcellation across a cohort, while dual regression approaches can be used to estimate session-specific parcellations from a group result.

Owing to these advantages and considering our strong experience in applying ICA to MEG data, we use spatial ICA as the technique for estimating a functional parcellation from the MEG data. That said, the methods that we go on to develop can be applied to any parcellation, regardless of its origin.

2. Estimating the network matrix: The parcellation stage yields a spatial map for each of the parcels. For each parcel, we can threshold the map to identify all the voxels that belong to that parcel. Then, we need to estimate a summary time course

that best represents the activity of the time courses of all the voxels belonging to that parcel. In fMRI, we might use the component time course as the summary time course (which corresponds to performing a multiple regression of the parcel maps from the concatenated fMRI data). However, in MEG we have sacrificed the high temporal resolution of the MEG data as part of our ICA procedure and so the IC time courses associated with each spatial IC will also have reduced temporal resolution. We would like a method for estimating a summary time course that regains the high temporal resolution of MEG.

In general, the summary time course of a parcel can be described as a weighted sum of all the voxel time courses belonging to that parcel. One option is to assign equal weighting to each voxel, effectively calculating the mean time course. However, if we consider a set of source-reconstructed voxel time series, each voxel has a sign ambiguity because we do not know the true orientation of the neuronal current dipoles (we introduced this idea in section 2.5.3.4). As such, the mean of the voxel time series is ill-defined. This was not an issue in our previous analyses, where we considered the envelope of each voxel (which is not affected by the sign ambiguity).

One alternative solution is to use the first principle component (PC) time course of the set of voxels as a parcel time course. This is equivalent to finding the set of weights that explains the maximum amount of variance of the set of voxels associated with a parcel. The first PC is unaffected by the dipole sign ambiguity. However, it may overly weight voxels with high variance due to non-neuronal artefacts.

Regardless of the technique used, a set of parcel time courses is produced. We define the network matrix as the all-to-all functional connectivity matrix. However, as we introduced in chapter 4, there are a range of measures that are used to describe FC.

In this chapter, we use two network matrices: the all-to-all *correlation matrix* and the all-to-all *covariance matrix* of the low-pass filtered envelopes of the parcel time courses. The standard correlation matrix is of interest as it links in with our seed-based correlation analyses (refer to section 4.4.2). The covariance matrix is of interest because we can use it to estimate the partial correlation matrix (see below) [95].

3. Analysing the network matrix: In the parcellation and network analysis stages, we have reduced the 4-dimensional MEG data into a single 2-dimensional matrix. At this stage, we can bring to bear a range of network analysis techniques that allow us to investigate the underlying structure of resting-state functional connectivity.

Such techniques include converting the network matrix into a graph, which is an abstract representation of the network consisting of nodes (parcels) and edges (functional connections). In this format, we can estimate graph-theoretical measures, such as degree (the number of connections that a node has) or betweenness (the number of times an edge of the graph is used as part of the set of shortest paths between all nodes) [126]. Each graph-theoretical measure attempts to capture a single feature of of a node or edge in the network graph with a scalar value. However, graph-theoretical measures are only as valid as the network matrices used to estimate them. As such, this chapter does not utilise graph-theoretical analyses but focuses the earlier stages of the network analysis pipeline.

In our analysis, we focus on one type of network analysis, namely *partial correlation*. Partial correlation is a technique for stripping away correlations between regions that are present due to an indirect interaction via other regions [95] (refer to section 1.5.1 for a discussion on the difference between *direct* and *indirect* FC). In principle, the partial correlation matrix is a measure of direct interactions between regions. This is of interest as it moves us one step closer to understanding how resting-state networks are causally structured and in particular what the consequences of network dysfunction might be.

7.2.2 Partial correlation analysis for MEG

Given a set of P parcels, each with a summary time course (estimated from source-reconstructed MEG data), we define standard FC as the correlation between the low-frequency ($<$0.5Hz) oscillations of the band-limited neural activity. Throughout this chapter, when we talk about correlation (both standard and partial) between parcels, we mean correlation between low-pass filtered envelopes of the band-limited

summary time courses.

For a pair of parcels, the standard correlation score (i.e. the Pearson correlation coefficient) is a measure of the linear dependency due to direct interactions between the two parcels and indirect interactions mediated by the other parcels. The partial correlation can be thought of as the correlation between the two parcels, after they have both been orthogonalised to the other $P-2$ parcels. Considering partial correlations in this way is conceptually useful. However, we do not actually estimate the partial correlation matrix this way. Instead, as we shall see, it is computed by inverting the covariance matrix.

7.2.2.1 Estimating partial correlations via the regularised inverse covariance matrix

The most popular technique for estimating partial correlations is via the inverse covariance method (ICOV). If we know the true covariance matrix, $\mathbf{\Sigma}$, of our parcellated data we can calculate the partial correlation matrix, $\mathbf{R}$, through the following relationship (taken from [95]):

$$R_{i,j} = -\frac{\Sigma^{-1}_{i,j}}{\Sigma^{-1}_{i,i}\Sigma^{-1}_{j,j}} \tag{7.1}$$

Marrelec et al. (2005) demonstrated that non-zero partial correlation scores will only exist if there are direct connections between network nodes. They did this by relating the data to a structural equation model (SEM), showing that connections (or non-zero elements) in the structural matrix correspond to non-zero elements in the precision matrix [93]. Critically, there is one exception to this finding. Non-zero partial correlations can occur between two nodes (A and B) that share no direct connections when both nodes act as inputs to a third node, C. In this case, the partial correlation analysis will actually infer a direct anti-correlation between A and B [124]. This is a consequence of Berkson's paradox, in which two independent variables can become conditionally dependent, given that at least one of them happens [133]. In this instance, the time courses of nodes A and B are independent. However, when we

estimate the partial correlation between A and B by conditioning out the time course of C (which is by definition a combination of A and B), we introduce a conditional dependency between A and B.

In general, we do not have access to the true covariance and precision matrices, $\mathbf{\Sigma}$ and $\mathbf{\Sigma}^{-1}$. However, we do have an estimate of the sample covariance matrix $\mathbf{\Sigma}_{sample}$ and can estimate a standard sample precision matrix $\mathbf{\Sigma}^{-1}_{sample}$. The standard sample precision matrix will not have any zero elements and so will estimate non-zero partial correlations between all nodes. Instead, we can estimate a sparse precision matrix by using an inversion which includes a penalty for non-sparsity. Specifically, we compute a maximum-likelihood estimate of the precision matrix that includes a L1-norm penalty on the estimated precision matrix, using the cost function [10]:

$$\underset{\mathbf{\Sigma}^{-1}_{sparse}}{max}\ (log(det(\mathbf{\Sigma}^{-1}_{sparse})) - trace(\mathbf{\Sigma}_{sample}\mathbf{\Sigma}^{-1}_{sparse}) - \lambda\|\mathbf{\Sigma}^{-1}_{sparse}\|_1) \tag{7.2}$$

where $\|\mathbf{\Sigma}^{-1}_{sparse}\|_1$ is the L1-norm of $\mathbf{\Sigma}^{-1}_{sparse}$ and is equal to the sum of the absolute of all its elements. We refer to this method as *L1-regularised* inversion and use an open source Matlab implementation.[1] By adjusting the L1-norm regularisation parameter, λ, we can control the sparsity of the estimated precision matrix.

Critically it should be noted that if we use our sparse estimate of the precision matrix, $\mathbf{\Sigma}^{-1}_{sparse}$, to estimate the partial correlation values using equation 7.1, then zero-off-diagonal elements in $\mathbf{\Sigma}^{-1}_{sparse}$ at $[i, j]$ will generate zero-elements in the partial correlation, $\mathbf{R}$, corresponding to no direct interaction between the i^{th} and j^{th} parcels.

7.2.3 Accounting for signal leakage

The *L1-regularised* ICOV method is a widely-adopted technique for estimating a sparse partial correlation matrix [124]. However, when applying it to the envelopes of source-space MEG data, we must account for signal leakage between voxels (and consequently between parcels). We introduced signal leakage in section 2.6 and cur-

[1]The Matlab implementation can be found at `http://www.di.ens.fr/~mschmidt/Software/L1precision.html`.

rent solutions that account for spurious FC due to signal leakage in section 4.4.1. In particular, we highlighted orthogonalisation as an efficient method for removing the spurious, zero-lag contribution of signal leakage [29, 74]. Orthogonalisation can be used to remove the spurious contribution to envelope correlations between a pair of voxels due to signal leakage. In this section, we consider the limitations of this approach as a method for accounting for signal leakage between parcels, and present a modified orthogonalisation approach which both corrects for signal leakage between parcels **and** allows the subsequent application of the ICOV method for estimating the partial correlation matrix.

7.2.3.1 Correctly accounting for signal leakage between two parcels

The signal leakage between a pair of voxels can be removed by regressing one voxel time course from the other, prior to envelope estimation. When presented with two parcels, each with a summary time course (defined as the first principle component), one might regress one summary time course from the other before estimating the envelope correlation. However, this does not account completely for signal leakage.

As an example, consider two parcels, A and B, each of which has been defined by finding regions whose voxels have strongly covarying oscillatory envelopes. We want to estimate if there is any correlation in the envelopes of the summary time courses of A and B that is indicative of underlying FC between the two parcels. For each parcel, we can map the voxel time courses into principle components via a PCA. Within each parcel, these PC time courses are by definition uncorrelated. However, the PCs will have some covariation in their envelope oscillations (as this is what was used to cluster the voxels into parcels).

We know that the voxels (and equivalently the PCs) from parcel A will have leaked into our estimate of the voxels/PCs of parcel B and vice versa. If we only regress the largest PC of parcel A out of parcel B, then parcel B will still contain some leaked contributions from all the other PCs of parcel A. These leaked PCs may be present in the envelope of the source-space estimate of the first PC of B. If we then calculate envelope correlations between the first PCs of A and B, we may be

measuring correlations between the envelope of the first PC of A and the envelopes of the leaked lower PCs of A that are present in the first PC of B. This means that we could measure non-zero FC even when parcels A and B have no true covariation in envelope oscillations.

Following this line of reasoning, it is clear that in order to account for all signal leakage between the two parcels we must regress all principle components corresponding to one parcel out of the first PC of the second parcel. We can then estimate the correlation between the orthogonalised first PCs of the two parcels. Furthermore, we can orthogonalise the two parcels using only the p largest PCs (where p corresponds to the number of PCs required to explain some fraction (e.g. 95%) of the parcel variance). This corresponds to accounting for 95% of the signal leakage.

7.2.3.2 Correctly accounting for signal leakage between multiple parcels

With the framework above, we can account for the signal leakage between pairs of parcels. We could implement this in a pair-wise approach and estimate a signal-leakage-corrected envelope correlation for every parcel pair. Similarly, we could estimate the signal-leakage-corrected envelope covariance between every pair. However, this is not equivalent to estimating a signal-leakage-corrected envelope covariance matrix. Covariance matrices are Gramian matrices (i.e. the inner product of two sets of vectors) and are consequently positive-semidefinite. In order to estimate a valid covariance matrix, we must estimate a single set of parcel time courses in which all signal leakage between all parcels has been accounted for and **then** estimate the envelope covariance matrix. To do this we have developed a procedure for mutually orthogonalising multiple (greater than two) parcels which we present in the following section.

7.3 Materials and methods

7.3.1 Data acquisition and preprocessing

To demonstrate our parcellation and network analysis framework, we acquired 10 minutes of *eyes-closed* resting-state data and 10 minutes of *eyes-open* resting-state data (where the subject had to fixate on a cross presented approximately 40cm ahead of them) for nine subjects. Data were acquired on a *CTF* 275-channel whole-head system. These data are a subset of the *CTF* resting/active-state data set used throughout this thesis. For a full description of the data (including demographics, structural MRI acquisition etc.) refer to section 5.5.1.1.

We applied our standard preprocessing analysis (described in full in section 3.6) which included 3rd order gradiometer correction, ICA de-noising to remove cardiac, eye-blink and mains artefacts, and removal of noisy channels and periods of data. We band-pass filtered the data into the 4-30Hz range (in order to span the theta, alpha and beta oscillatory bands). We performed source reconstruction using an LCMV beamformer to estimate neural activity across a 6mm grid spanning the whole brain for use in the parcellation via spatial ICA. We also separately beamformed the data to a 10mm grid spanning the whole brain for use in the estimation of parcel time courses. 10mm resolution data was estimated because the mutual orthogonalisation procedure requires multiple PCA decompositions and is extremely slow for 6mm resolution data. For the purpose of the development and proof of concept presented in this chapter, 10mm resolution was more than sufficient.

7.3.2 Parcellating MEG data using spatial ICA

In an identical fashion to the temporal ICA analyses presented in chapters 4 and 5, for each session we estimated the envelopes of the band-limited beamformed data, applied *weights normalisation* and then performed a windowed-average down-sampling (using a window length of 2s). We performed 4mm (FWHM) spatial smoothing, spatial resampling, and de-meaning, followed by a temporal concatenation of all the sessions. Unlike previous analyses, we performed a **spatial** ICA decomposition to estimate

40 spatially independent component maps and associated time courses. However, in practical terms, this simply corresponds to transposing the concatenated data matrix and then applying **FastICA** in an identical manner as for temporal ICA.

We selected a model order of 40 for the following reason. In MEG, we have observed that spatial ICA tends to always give single region components. We have only noted the consistent emergence of multi-node components when using model orders below 10. However, model orders between 10 and 20 produce very large single node component maps that cover multiple RSN nodes. By using a model order of 40, we obtained single node ICs with spatial extents that approximately match those of individual RSN nodes (such as the anterior and posterior nodes of the frontoparietal networks shown in Figure 7.1).

The spatial ICA yielded 40 group-average spatially independent components. We performed a multiple regression of the session-specific component time course segments out of the session-specific voxel-level envelopes to produce 40 parcel maps per session. The spatial maps for each session were binarised into parcel masks by assigning each voxel to the parcel with the largest weighting for that voxel (in a winner-takes-all fashion).

7.3.3 Network matrix estimation

The procedure for estimating the set of mutually orthogonal parcel time courses for each session is described below. For each session of 10mm resolution, beamformed data, we:

1. Randomly selected an initial parcel.
2. Estimated the p largest principle components of the voxels within that parcel.
3. Regressed the p largest principle components out of the time course of every voxel that does not belong to the parcel.
4. Kept the first PC (of the p PCs estimated in step 2) as a summary time course for that parcel.

5. Randomly selected the next parcel.

6. Estimated the p largest principle components of the voxels (which have been orthogonalised to the first parcel) within that parcel.

7. Regressed the p largest principle components out of the time course of every voxel that does not belong to that parcel (nor any of the parcels that have already been orthogonalised).

8. Kept the first PC (of the p PCs estimated in step 6) as a summary time course for that parcel.

9. Repeated stages 5-8 until every parcel had a summary time course that was mutually orthogonal to all the other parcels.

We selected $p = 12$, as we found that this approximately accounted for 95% of the total signal leakage. At each stage of the multi-parcel orthogonalisation, this meant that the rank of our data was reduced by 12. The choice of p places a limit on the number of parcels that we can orthogonalise. *CTF* data have a rank of 275, limiting us to a maximum of $\frac{275}{p}$ orthogonal parcels. This is because each parcel needs its own 12-dimension subspace in order to be 95% orthogonal to all the other parcels. 22 parcels will span the entire 275-dimension space of the *CTF* data, with the consequence that we cannot define any further subspaces that are orthogonal to all the other parcels. As such, we selected 22 parcels that corresponded to known RSN nodes and discarded the remaining 18.

This procedure yielded 22 mutually orthogonal time courses for each session of data. We simultaneously estimated the largest PC of the 22 parcels without applying orthogonalisation so that we could demonstrate the efficacy of our signal leakage correction. We band-pass filtered each parcel time course into three frequency bands (4-8Hz, 8-13Hz and 13-30Hz) to generate 66 band-specific parcel time courses for each session. This allowed us to investigate the direct functional connectivity between parcels and across frequency bands, which is an insight unique to MEG.

We temporally concatenated the 66 orthogonalised parcel time courses from every session to obtain a single group set of parcel time courses. Concatenation is essential to impose the same sparsity constraint uniformly across every session of data in the *L1-regularised* estimation of the precision matrix. Using the group-concatenated time courses, we estimated the correlation and covariance matrix between the low-pass filtered (< 0.5Hz) envelopes of the summary time courses. We estimated similar correlation and covariance matrices for the non-orthogonalised parcel envelopes.

7.3.4 Partial correlation analysis

In order to estimate each partial correlation matrix (without and with multi-parcel orthogonalisation), we computed the sparse estimates of the precision matrices via the *L1-regularised* inversion (with a regularisation parameter $\lambda = 0.1$). We estimated the partial correlation matrices by inputting these sparse precision matrices into equation 7.1.

7.4 Results

7.4.1 Parcellations based on spatial ICA

Figure 7.1 shows the 22 parcels used in the network analysis. The parcels were manually selected from a set of 40 spatially independent components based on a manual matching of each parcel to a node of a known resting-state network. Parcels corresponding to nodes of the default mode and frontoparietal networks have been highlighted. Parcels that were discarded corresponded to white matter, cerebellum and grey matter regions not linked to the main RSNs.

7.4.2 Cross-frequency standard and partial correlation matrices without multi-parcel orthogonalisation

Figure 7.2 shows the results of our cross-frequency network analysis without accounting for spurious inter-parcel correlations due to signal leakage. Figure 7.2A is the

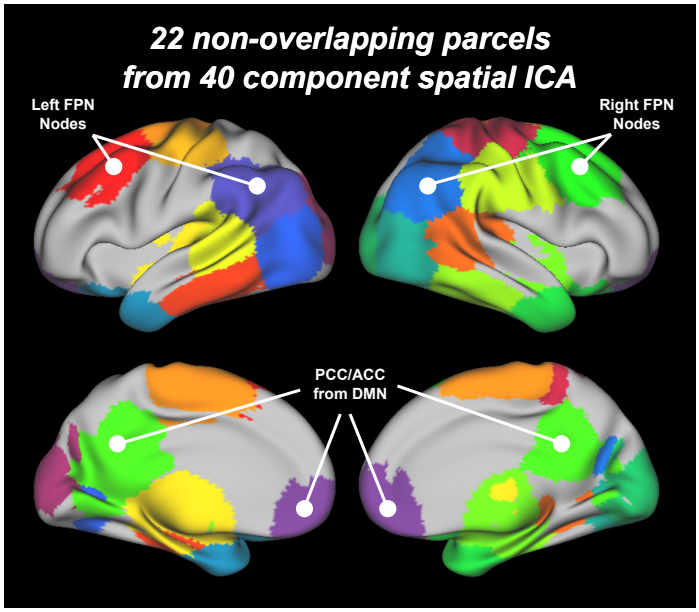


Figure 7.1: *The 22 non-overlapping parcels masks used in the network analysis, estimated using a whole-brain, 40-component spatial ICA. Parcels were selected based on their correspondence to nodes of the following resting-state networks: left and right frontoparietal networks; sensorimotor network; default mode network; bilateral temporal lobes; temporoparietal junction; inferior temporal gyri; visual network. We have highlighted parcels corresponding to: the PCC (green) and ACC (purple) nodes of the DMN; the anterior (red) and posterior (blue) nodes of the left frontoparietal network (FPN); the anterior (green) and posterior (blue) nodes of the right frontoparietal network (FPN). All parcels have been projected onto an inflated cortical surface for visualisation.*

standard correlation matrix for the low-pass filtered envelopes of the first PCs of each parcel (band-pass filtered into the theta, alpha and beta bands). Figure 7.2B is the equivalent partial correlation matrix. It is worth noting that the sparse precision matrix estimation has removed a large number of connections. However, the surviving partial correlations are dominated by neighbouring parcels as exemplified by Figures 7.2C and D. Using the posterior node of the right frontoparietal network as a seed, we have visualised the standard (C) and partial (D) correlations with the remaining parcels (in the beta band). In both cases, the strongest FC is measured between the seed parcel and the neighbouring motor cortex parcels.

7.4.3 Cross-frequency standard and partial correlation matrices with multi-parcel orthogonalisation

Figure 7.3 shows the results of our cross-frequency network analysis after applying our multi-parcel orthogonalisation to remove spurious zero-lag inter-parcel correlations

due to signal leakage. Figure 7.3A is the standard correlation matrix for the low-pass filtered envelopes of the first PCs of each parcel (band-pass filtered into the theta, alpha and beta bands). Figure 7.3B is the equivalent partial correlation matrix. Again, the regularised precision matrix estimation has pruned away the correlations due to indirect interactions. Importantly, in this analysis, the strongest correlations are no longer between neighbouring parcels. We have again visualised the envelope correlations of the beta band parcel time courses (C. standard and D. partial) between the posterior node of the right FPN and all other parcels. The correlation values are taken from the elements of each network matrix marked with a black box. The strongest correlations, for both the full and partial correlations, are now between the seed and the anterior node of the right FPN, demonstrating that the multi-parcel orthogonalisation has stripped out the signal leakage component of FC which is mostly from neighbouring regions.

7.5 Discussion

Spatial ICA for parcellating MEG data: We have demonstrated that spatial ICA is capable of parcellating MEG data into a set of functionally relevant regions. The parcellation from a 40 component spatial ICA yields a set of single region non-overlapping parcels. Furthermore, these parcels have a sensible correspondence to nodes of RSNs. Finally, the parcellation generates approximately bilateral parcel pairs. The spatial ICA has no built-in prior to encourage symmetry. The symmetrical pairing can only reflect the inherent symmetry of the brain's natural parcellation. The spatial ICA used here represents an initial attempt but is by no means the final word in MEG-derived parcellations. As with other ICAs, selecting the model order is a key question. In particular, the following two questions are of key interest. Firstly, which model order will give the best correspondence to individual RSN nodes? Secondly, which model order will give the best delineation of anatomically distinct regions? We have found very preliminary evidence that high model orders (≥ 80) may actually provide a reasonable delineation between grey matter, white matter

and the cerebellum and think this warrants future investigation. Finally, although we have utilised a MEG-derived parcellation in this analysis, alternative parcellations, such as those derived from a high-dimensional spatial ICA on fMRI data, may offer complementary and possibly superior results.

Multi-parcel orthogonalisation: We have provided strong evidence that standard and partial correlation based network analyses are just as susceptible to spurious FC as the voxel-wise seed-based correlation analyses presented in section 4.4.2. However, all-to-all network analysis methods (such as partial correlation) cannot utilise the standard pairwise orthogonalisation correction [29,74]. Instead, we have presented an alternative technique for estimating a set of mutually orthogonal parcel time courses. We demonstrated that this technique is able to remove the dominant spurious correlations between neighbouring parcels. However, our multi-parcel orthogonalisation has a major drawback. Our technique is limited by the inherently low spatial rank of MEG data. At each stage of the orthogonalisation, we have to reduce the rank of the data by some number of dimensions (in this analysis, 12). This limits the number of parcels that can be analysed robustly. Using our framework, there are three ways to respect the dimensionality limit. 1.) Only consider a small number of parcels - this is what we did in our analysis. 2.) Erode the parcels to the point that a small enough number of PCs is sufficient to model all the signal leakage. In the limit, we could consider 275 single voxel parcels (when analysing *CTF* data). 3.) Only use a small number of PCs for each parcel in the multi-parcel orthogonalisation. This corresponds to weakening our orthogonalisation protection against signal leakage and in the limit would correspond to no orthogonalisation. In general, the multi-parcel orthogonalisation method cannot provide both rigorous protection against signal leakage **and** whole-brain coverage. Depending on the analysis in question, the user must prioritise one of these factors when implementing the analysis.

The upper bound on the number of parcels that can be mutually orthogonalised is a critical limitation in the case of partial correlation analyses. Partial correlation is a robust technique for estimating the strength of direct connections in a network

when every node of the network is being modelled. However, in the presence of one or more missing nodes, we can no longer state that a non-zero partial correlation score proves the existence of a direct interaction. Whole-brain network analyses attempt to capture all of the brain's activity. However, our use of a subset of parcels exposes our analysis to the missing node confound.

One question that we have not addressed is whether the order in which the parcels are orthogonalised matters. In principle, we do not foresee a reason why the ordering should significantly change the output but we do concede that different orderings will lead to different numerical values of FC. This is even apparent in pairwise orthogonalisation where the solution is to perform both possible orthogonalisations and average the two results. However, the combinatorial explosion of possible orderings (in this instance there are $22! = 1.2 \times 10^{21}$ possible orders) makes this impossible. In our analysis, to mitigate any ordering bias, we randomised the order of orthogonalisation for each session of data. In future applications of this technique we would advocate performing multiple runs to establish whether any interesting FC is stable across several different orderings.

Partial correlation: In this analysis, we compared partial correlation to standard correlation. Most importantly, we have demonstrated that partial correlation does remove many indirect connections from the network matrix. While the missing node confound means that we cannot state with complete certainty that the non-zero partial correlation indicates the existence of a direct interaction between parcels/frequency bands, we can say that zero partial correlation scores indicate an absence of direct interaction between parcels/frequency bands. This does assume that the sparsity in our estimate of the precision matrix reflects genuine sparsity in the true network matrix and is not an artefact of the choice of the regularisation parameter, λ. Exploring the effect of changing λ is beyond the scope of this preliminary investigation into MEG network analysis but will be a critical question to address in future work.

Assuming our choice of λ is sensible, we can see that the *L1-regularised* precision

matrix forced many elements of the signal-leakage-corrected partial correlation matrix to zero. Of particular interest are the non-zero elements in the off-diagonal blocks of the matrix, which correspond to cross-frequency FC. Our analysis indicates that the dominant cross-frequency FC between envelope time courses exists within each parcel. However, we also found evidence of envelope correlations between different frequency bands of spatially separate parcels. Although these FC scores are extremely weak, the fact that they have survived the sparsity penalty suggests they are genuine. Furthermore, within-frequency band elements of the partial correlation matrix have been forced to zero, suggesting that our non-zero cross-frequency/cross-parcel partial correlation scores are not an artefact of under-penalising non-sparsity in our estimate of the precision matrix.

7.6 Summary

In this chapter, we presented an initial framework for performing a parcellation and subsequent network analysis on MEG data. We showed that spatial ICA can be used to estimate a data-driven parcellation of the brain. We placed particular emphasis on implementing a novel signal leakage correction procedure that is compatible with multi-parcel network modelling techniques like ICOV-based partial correlation. Accounting for signal leakage is the single biggest challenge in any MEG FC analysis and network modelling approaches are no exception. We demonstrated that the signal leakage confound dominates the estimates of both standard and partial correlation matrices, and showed that our multi-parcel orthogonalisation successfully removes this spurious FC from both correlation matrices. Finally, we used our MEG network analysis pipeline to look at FC between slow envelope oscillations of different frequency bands both within and between parcels.

Much like the MEG-adapted dual regression technique presented in chapter 5, our network analysis framework does not require a MEG-derived parcellation but can use any parcellation. This makes it an excellent tool for cross-modality validation and investigation into the structure of functional networks.

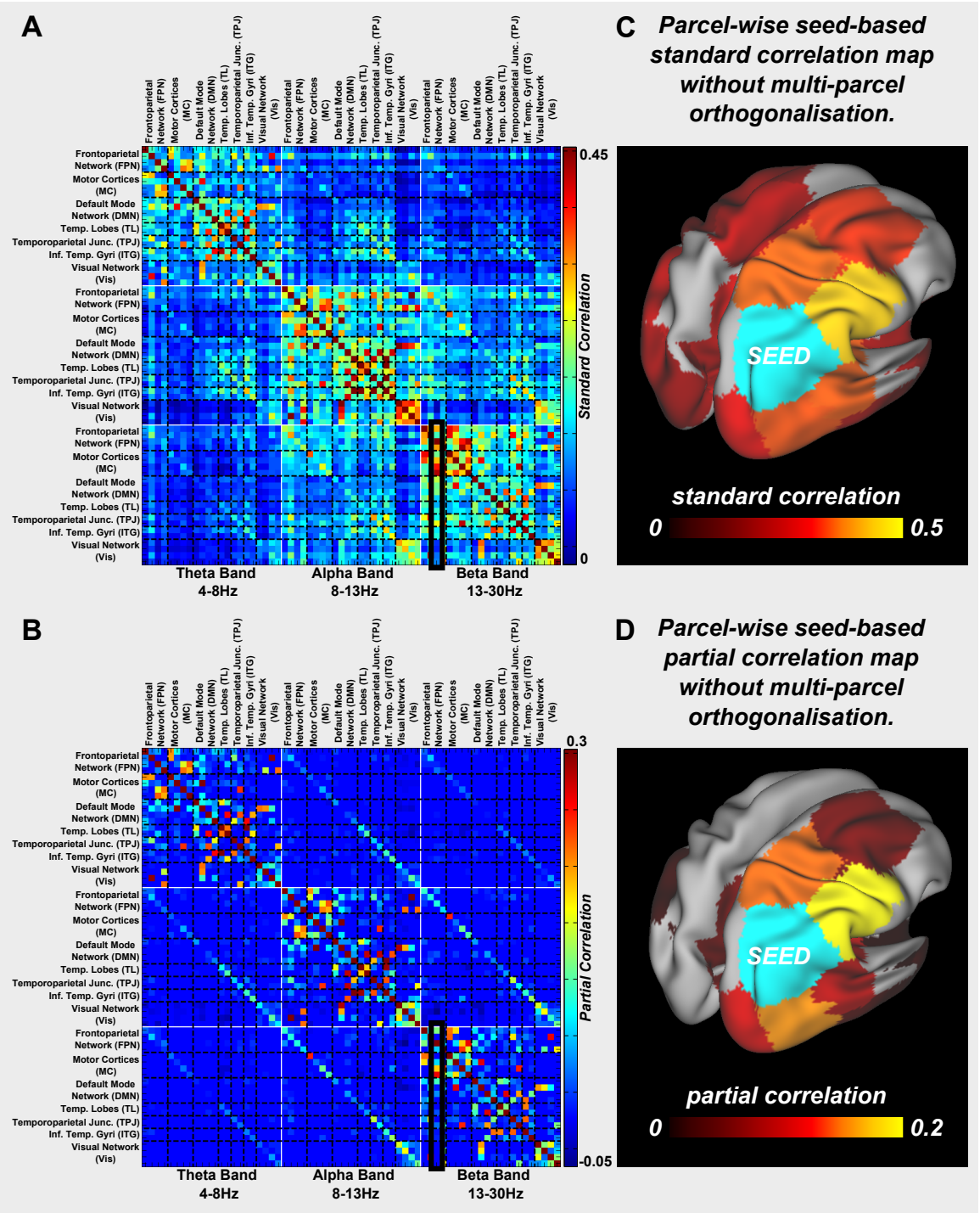


Figure 7.2: *Results of cross-frequency, multi-parcel network analyses without any multi-parcel orthogonalisation to account for signal leakage. A. The matrix of Pearson correlation coefficients between the low-pass filtered envelopes of the band-limited parcel time courses (concatenated over subjects). B. The equivalent partial correlation matrix, estimated via the inverse-covariance matrix method. C. Parcel-wise seed-based correlation map using the posterior node of the right frontoparietal network as a seed (shown in blue). D. Parcel-wise seed-based partial correlation map using the posterior node of the right frontoparietal network as a seed (shown in blue). The correlation values are taken from the elements of each network matrix marked with a black box. In both cases note that neighbouring parcels show the strongest correlation and partial correlation.*

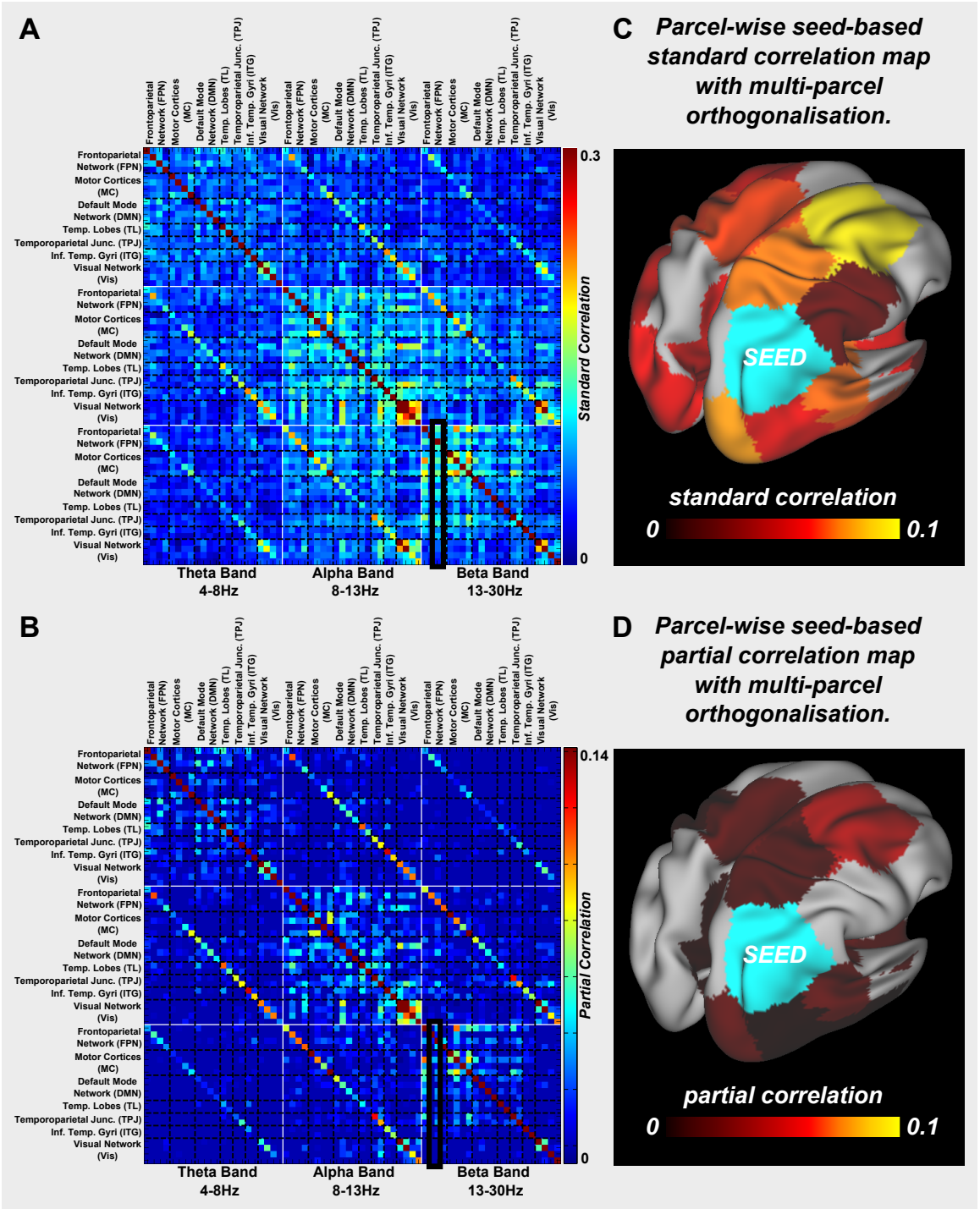


Figure 7.3: *Results of cross-frequency, multi-parcel network analyses after applying multi-parcel orthogonalisation to account for signal leakage. A. The matrix of Pearson correlation coefficients between the low-pass filtered envelopes of the band-limited parcel time courses (concatenated over subjects). B. The equivalent partial correlation matrix, estimated via the inverse-covariance matrix method. C. Parcel-wise seed-based correlation map using the posterior node of the right frontoparietal network as a seed (shown in blue). D. Parcel-wise seed-based partial correlation map using the posterior node of the right frontoparietal network as a seed (shown in blue). In both cases note that the anterior node of the right frontoparietal network is now showing the strongest direct and indirect FC.*

Chapter 8

Conclusions and future work

This thesis is a collection of methods for performing functional connectivity analyses in MEG, with a particular focus on resting-state data. We have presented novel insights at every stage of the analysis, from data preprocessing, to MEG source reconstruction, to methods for estimating connectivity, to group statistical methods for comparing functional networks and exploring the underlying network structure.

AD has been identified as one of one of the greatest emerging healthcare challenges. The motivation behind developing the methods presented in this thesis is to facilitate the exploration of the effect of the APOE-ϵ4 genotype, a proven risk factor for AD, on resting-state brain functional connectivity using the direct, whole-brain, electrophysiological measure of neural activity afforded by MEG. We hope that better understanding of the causal mechanism by which APOE-ϵ4 promotes the development of AD will lead to a more complete understanding of the disease and the development of effective, early interventions to avoid AD-related dementia.

8.1 Summary of the novel contributions of this thesis

In **chapter 3**, we presented a preprocessing pipeline for de-noising MEG data and performing source reconstruction using a beamformer. We made significant advances in developing a robust framework for processing MEG data acquired on an *Elekta Neuromag* system. This included an iterative manual procedure for identifying specific

scanner artefacts that are amplified and propagated by the application of signal space separation via *MaxFilter*^{TM} [50, 127]. We presented minimum eigenvalue normalisation as a novel technique that enables the de-noising of multi-sensor-type data using ICA. We also demonstrated that minimum eigenvalue normalisation is essential for combining information from multiple sensor types during beamforming. Finally, we demonstrated that our beamformer implementation can robustly source-reconstruct *MaxFiltered* data by performing a reduced-rank inversion of the data covariance matrix. We combined all of these modifications into a single preprocessing pipeline that is compatible with both resting-state and task-positive MEG data, acquired on either *Elekta Neuromag* or *CTF* scanners. This preprocessing pipeline underpinned a range of subsequent source-space functional connectivity analyses.

In **chapter 4**, we considered how best to measure functional connectivity in source-space MEG data and how best to image resting-state networks. We demonstrated that correlations between the slow oscillations in the envelope of band-limited neural activity provide a robust measure of FC with excellent correspondence to fMRI. We presented an investigation into the optimum *carrier* and *co-variation* frequencies for measuring envelope correlations in the presence of signal leakage, and showed that FC is best measured in the slow envelope oscillations (0.05-0.5Hz) of the 8-20Hz band of neural activity.

We established a seed-based correlation approach for imaging single RSNs in MEG. While not novel in itself, this seed-based correlation framework is now being used to generate *carrier/co-variation* frequency specific envelope correlation matrices which are being used to validate the accuracy of a range of biophysical models (in a similar way to the work of Cabral et al. (2011) [34]).

We developed a framework for decomposing the concatenated envelopes of source-reconstructed MEG data using group ICA. Using this framework, we reproduced the critical finding of Brookes et al. (2011) that demonstrates that RSNs have an electrophysiological basis and can be imaged in MEG without prior information from other modalities. Furthermore, we demonstrated that our analysis can detect similar RSNs in *Elekta Neuromag* data.

In **chapter 5**, we presented a statistical framework for comparing RSNs between different populations. As part of the development, we showed how *weights normalisation* introduces a major confound in analyses where multiple sessions are separately beamformed. We subsequently presented MEG-adapted dual regression (MADR) as a technique for estimating differences between resting-state activity of sub-group populations given any group-level spatial basis set. We used this analysis to demonstrate a significant reduction in the variance of oscillatory power associated with both the left and right frontoparietal networks in carriers of the APOE-ϵ4 gene compared with normal controls (APOE-ϵ3 homozygotes). This finding is a major result and represents a starting point for a range of future analyses aimed at discovering how the APOE-ϵ4 gene modulates resting-state brain function at the electrophysiological level, resulting in a predisposition toward AD. This includes the work by Heise (2013) who uses the MADR framework to perform a thorough investigation into the effect of the APOE genotype on resting-state FC in MEG (refer to Heise's D.Phil thesis [71], currently under examination).

In **chapter 6**, we outlined a framework for estimating functionally-connected networks in task-positive MEG data. We presented a statistical analysis, using a mixed effects GLM, capable of separating task-relevant networks from task-irrelevant components. We showed that this framework offers substantial advantages over the ICA or the GLM (or similar mass univariate statistics) when used in isolation.

In **chapter 7**, we presented an initial investigation into resting-state MEG using a functional parcellation combined with a partial correlation network analysis. We showed that spatial ICA can decompose MEG data into a useful, data-driven parcellation. We developed a novel approach for estimating a mutually orthogonal set of parcel time courses, given any general parcellation. We applied this multi-parcel orthogonalisation technique to our MEG-derived parcellation, and applied a partial correlation analysis to the result to show evidence of cross-frequency envelope correlations within and between parcels.

8.2 Alternative applications of MEG functional connectivity methods

A substantial part of this D.Phil has been spent incorporating the analysis pipelines presented in chapters 3, 4, 5, and 6 into **OHBA's Software Library** (**OSL**), an open-source MEG analysis toolbox. The resting-state functional connectivity analyses in **OSL** have been used to investigate the interaction between APOE-ϵ4, functional connectivity in MEG and Alzheimer's disease. However, they can be applied to a range of other clinically and neuro-scientifically interesting applications

For example, the ZNF804A gene has been associated with schizophrenia [105]. In fMRI, there is growing evidence that this genotype leads to abnormal FC [51, 113]. Furthermore, abnormal RSN function has been implicated in schizophrenia pathology, such as that associated with the salience network [109, 132]. Investigating the role of ZNF804A in healthy adults and schizophrenics using MEG would be an alternative application of the methods developed in this thesis.

Another example is the interaction of dopamine on learning processes. In animal models, dopamine (DA) transmission has been shown to influence successful learning by modulating low-frequency (theta-band) synchronisation between the hippocampus and the prefrontal cortex [18, 39, 61]. Natural variation in the cortical availability of DA (for instance, through differences in the COMT-genotype), or its manipulation through the administration of dopaminergic drugs may also influence such synchronisation in humans. While studies using fMRI [130] have demonstrated prefrontal connectivity changes as a function of COMT genotype, the analysis of resting-state networks in MEG can provide a more direct test of the hypothesis that DA influences synchronisation in a particular frequency band.

8.3 Future directions

In this section, we consider the future directions in which we could proceed based on the corpus of work presented in this thesis.

8.3.1 Modelling non-stationarity in FC

All the techniques that we have used to measure functional connectivity assume temporal stationarity over the entirety of each session of data. However, much evidence is emerging that shows that this assumption does not reflect genuine brain function [76]. For example, de Pasquale et al. (2012) have shown with MEG that the PCC node of the DMN dynamically integrates with other resting-state networks in a unique fashion not seen in other RSNs [44]. Allen et al. (2012) have shown that resting-state fMRI data can be broken down into a set of repeating but varied patterns of functional connectivity [3]. Most interestingly, Baker et al. have shown that consistent modes of functional connectivity can be extracted from resting-state MEG data. These modes correspond to the stationary RSNs seen in fMRI and MEG but are switching at time scales of approximately a hundred milliseconds [8].

The stationary RSNs that we detect when using tools like ICA and seed-based correlation correspond to temporal averages of these highly dynamic functional connectivity processes that have been discovered recently. While these temporal average RSNs are extremely informative, for example as biomarkers for disease, exploring the temporal dynamics of neuronal functional connectivity presents the best opportunity for furthering our understanding of human brain function. MEG is ideally suited to contribute to this goal as it offers the best whole-brain spatial resolution of any modality (with the exception of concurrent EEG/fMRI) capable of recording neuronal activity at the time scales of cognitive processing [62].

8.3.2 Multi-modal analysis and the Human Connectome Project

While substantial advances have been made in performing functional connectivity analyses on source space MEG data, both in this thesis and in other studies [25, 26, 30, 43, 44, 74, 90], MEG's spatial resolution is still much poorer than that of fMRI. Developing analysis tools that can use the excellent spatial resolution of fMRI to inform the exploration of the rich temporal dynamics of neuronal activity in MEG will be essential for gleaning as much useful information as possible from MEG data. We

presented several methods that can utilise fMRI-derived spatial basis sets to explore functional connectivity in MEG (refer to chapters 5 and 7). However, such methods only represent a starting point. Ideally, we would like to incorporate the spatial information of fMRI into our entire analysis, including source reconstruction [121].

Such techniques could be of immense use with the emerging availability of the multi-modal neuroimaging data acquired as part of the *Human Connectome Project* (HCP). The HCP includes high quality fMRI, diffusion tractography, and MEG data from a single cohort of 50 twin pairs [85, 121]. Having a framework for combining information from these three modalities will maximise the potential insight that this data set can provide.

8.4 Concluding remarks

The exploration of functional connectivity in the human brain represents a major step in moving from perceiving the brain as a static neural assembly to a dynamic system. With such a perspective, we can make major advances in understanding what drives healthy brain function and accordingly discover and prevent the causes of various neurological diseases. In recent years, MEG research into functional connectivity has moved from replicating the insights of fMRI to becoming an essential research tool. This thesis represents significant work in developing the tools that have helped to bridge the gap between MEG and fMRI, and will enable us to uncover new insights into human brain function.

Appendix A

List of journal and conference publications

D. E. Astle, H. Luckhoo, M. W. Woolrich, B. Kuo, A. C. Nobre, and G. Scerif. *The neural dynamics of fronto-parietal networks in childhood revealed using magnetoencephalography.* Cerebral Cortex, (Under Review), 2013.

M. J. Brookes, M. W. Woolrich, H. Luckhoo, D. Price, J. R. Hale, M. C. Stephenson, G. R. Barnes, S. M. Smith, and P. G. Morris. *Investigating the electrophysiological basis of resting state networks using magnetoencephalography.* Proceedings of the National Academy of Sciences of the United States of America, 108(40):16783-16788, Sept. 2011.

M. J. Brookes, E. B. Liddle, J. R. Hale, M. W. Woolrich, H. Luckhoo, P. F. Liddle, and P. G. Morris. *Task induced modulation of neural oscillations in electrophysiological brain networks.* NeuroImage, 63(4):1918-30, Dec. 2012.

H. Luckhoo, J. R. Hale, M. G. Stokes, A. C. Nobre, P. G. Morris, M. J. Brookes, and M. W. Woolrich. *Inferring task-related networks using independent component analysis in magnetoencephalography.* NeuroImage, 62:530-541, May 2012.

H. Luckhoo, M. J. Brookes, V. Heise, C. E. Mackay, K. Ebmeier, P. G. Morris, and M. W. Woolrich. *Extracting resting state networks from Elekta Neuromag MEG data using independent component analysis.* In 18th Annual Meeting of the Organization for Human Brain Mapping, 2012.

H. Luckhoo, M. J. Brookes, V. Heise, C. E. Mackay, K. Ebmeier, P. G. Morris, and M. W. Woolrich. *Investigating the frequency composition of resting state networks in MEG using ICA*. In 18th Annual Meeting of the Organization for Human Brain Mapping, 2012.

H. Luckhoo, M. J. Brookes, S. Palazzo Corner, A. Baker, and M. W. Woolrich. *Measuring cross-frequency functional connectivity in resting-state MEG*. In 19th Annual Meeting of the Organization for Human Brain Mapping, 2012.

H. Mohseni, M. W. Woolrich, M. Kringelbach, H. Luckhoo, P. Probert Smith, and T. Z. Aziz. *Fusion of Magnetometer and Gradiometer Sensors of MEG in the Presence of Multiplicative Error*. IEEE Transactions on Biomedical Engineering, 59(7):1951-61, July 2012.

M. W. Woolrich, A. Baker, H. Luckhoo, H. Mohseni, G. Barnes, M. J. Brookes, and I. Rezek. *Dynamic State Allocation for MEG Source Reconstruction*. NeuroImage, 77:77-92, July 2013.

Appendix B

List of Abbreviations

B.1 Glossary of Abbreviations

- **ACC**: Anterior Cingulate Cortex.
- **AD**: Alzheimer's Disease.
- **APOE**: Apolipoprotein E.
- **AU**: Arbitrary Units.
- **BEM**: Boundary Element Model.
- **BLP**: Band Limited Power.
- **BOLD**: Blood Oxygen Level Dependent.
- **COPE**: Contrast of Parameter Estimates.
- **DAN**: Dorsal Attention Network.
- **DMN**: Default Mode Network.
- **ECG**: Electrocardiogram.
- **EEG**: Electroencephalography.
- **EPSP**: Excitatory Post-synaptic Potential.
- **ERS/ERD**: Event-Related Synchronisation/Desynchronisation.
- **FC**: Functional Connectivity.
- **FCN**: Functionally-Connected Network.
- **FDR**: False Discovery Rate.
- **FIL**: Functional Imaging Laboratory, University College, London.
- **fMRI**: Functional Magnetic Resonance Imaging.
- **FMRIB**: Centre for Functional Magnetic Resonance Imaging of the Brain, University of Oxford.
- **FPN**: Frontoparietal Network.
- **FSL**: FMRIB's Software Library.
- **FWHM**: Full Width at Half Maximum.
- **GLM**: General Linear Model.
- **HCP**: Human Connectome Project.
- **HPI**: Head Position Indicator.

- **HRF**: Haemodynamic Response Function.
- **ICA**: Independent Component Analysis.
- **ICOV**: Inverse Covariance Matrix Method.
- **IPSP**: Inhibitory Post-synaptic Potential.
- **LCMV**: Linearly Constrained Minimum Variance.
- **MADR**: MEG-Adapted Dual Regression.
- **MC**: Motor Cortex/Cortices.
- **MCW**: Maximally Correlated Window.
- **MEG**: Magnetoencephalography.
- **MNE**: Minimum Norm Estimate.
- **MNI**: Montreal Neurological Institute.
- **MSP**: Multiple Sparse Priors.
- **MSR**: Magnetically Shielded Room.
- **OHBA**: Oxford Centre for Human Brain Activity.
- **OSL**: OHBA's Software Library.
- **PC**: Principle Component.
- **PCA**: Principle Component Analysis.
- **PCC**: Posterior Cingulate Cortex.
- **PSSC**: Primary Somatosensory Cortices.
- **PET**: Positron Emission Tomography.
- **ROI**: Region of Interest.
- **RSN**: Resting-State Network.
- **SAM**: Synthetic Aperture Magnetometry.
- **SEM**: Standard Error of the Mean.
- **SEM**: Structural Equation Model.
- **SMN**: Sensorimotor Network.
- **SPL**: Superior Parietal Lobule.
- **SQUID**: Superconducting Quantum Interference Device.
- **SSN**: Somatosensory Network.
- **SSS**: Signal Space Separation.
- **SVD**: Singular Value Decomposition.
- **TFCE**: Threshold-Free Cluster Enhancement.
- **TPJ**: Temporoparietal Junction.
- **WM**: Working Memory.
- **WRT**: With Respect To.

Bibliography

[1] A. Abou-Elseoud, T. Starck, J. Remes, J. Nikkinen, O. Tervonen, and V. Kiviniemi. The effect of model order selection in group PICA. *Human Brain Mapping*, 31(8):1207–16, Aug. 2010.

[2] R. Adalbert, J. Gilley, and M. P. Coleman. Aβ, tau and ApoE4 in Alzheimer's disease: the axonal connection. *Trends in Molecular Medicine*, 13(4):135–142, Apr. 2007.

[3] E. A. Allen, E. Damaraju, S. M. Plis, E. B. Erhardt, T. Eichele, and V. D. Calhoun. Tracking Whole-Brain Connectivity Dynamics in the Resting State. *Cerebral Cortex*, pages 1–14, Nov. 2012.

[4] D. E. Astle, H. Luckhoo, M. W. Woolrich, B. Kuo, A. C. Nobre, and G. Scerif. The neural dynamics of fronto-parietal networks in childhood revealed using magnetoencephalography. *Cerebral Cortex*, (Under Review), 2013.

[5] N. Axmacher, M. M. Henseler, O. Jensen, I. Weinreich, C. E. Elger, and J. Fell. Cross-frequency coupling supports multi-item working memory in the human hippocampus. *Proceedings of the National Academy of Sciences of the United States of America*, 107(7):3228–33, Feb. 2010.

[6] A. Baddeley. Working Memory. *Science*, 255(5044):556–559, July 1992.

[7] S. Baillet, J. C. Mosher, and R. M. Leahy. Electromagnetic Brain Mapping. *IEEE Signal Processing Magazine*, 18(6):14–30, 2001.

[8] A. Baker, M. J. Brookes, H. Luckhoo, I. Rezek, P. Brodersen, S. Palazzo Corner, P. Probert Smith, and M. W. Woolrich. Inferring Transiently Synchronising

Networks using a Hidden Markov Model. In *19th Annual Meeting of the Organization for Human Brain Mapping*, 2013.

[9] A. Baker, M. J. Brookes, P. Probert Smith, and M. W. Woolrich. Investigating the temporal dynamics of resting state connectivity with MEG. *18th Annual Meeting of the Organization for Human Brain Mapping*, 2012.

[10] O. Banerjee, L. Ghaoui, A. D'Aspremont, and G. Natsoulis. Convex optimization techniques for fitting sparse Gaussian graphical models. In *Proceedings of the 23rd International Conference on Machine Learning*, pages 89–96, 2006.

[11] G. R. Barnes and A. Hillebrand. Statistical flattening of MEG beamformer images. *Human Brain Mapping*, 18(1):1–12, Jan. 2003.

[12] J. T. Becker, M. Fabrizio, G. Sudre, A. Haridis, T. Ambrose, H. J. Aizenstein, W. Eddy, O. L. Lopez, D. A. Wolk, L. Parkkonen, and A. Bagic. Potential utility of resting-state magnetoencephalography as a biomarker of CNS abnormality in HIV disease. *Journal of Neuroscience Methods*, 206(2):176–82, Jan. 2012.

[13] C. F. Beckmann, M. DeLuca, J. T. Devlin, and S. M. Smith. Investigations into resting-state connectivity using independent component analysis. *Philosophical transactions of the Royal Society of London.*, 360(1457):1001–13, May 2005.

[14] C. F. Beckmann, C. E. Mackay, N. Filippini, and S. M. Smith. Group comparison of resting-state FMRI data using multi-subject ICA and dual regression. In *Organisation for Human Brain Mapping*, July 2009.

[15] C. F. Beckmann and S. M. Smith. Probabilistic independent component analysis for functional magnetic resonance imaging. *IEEE Transactions on Medical Imaging*, 23(2):137–52, Feb. 2004.

[16] A. J. Bell and T. J. Sejnowski. An information-maximization approach to blind separation and blind deconvolution. *Neural Computation*, 7(6):1129–59, Nov. 1995.

[17] A. Belouchrani, K. Abed-Meraim, J.-F. Cardoso, and E. Moulines. A blind source separation technique using second-order statistics. *IEEE Transactions on Signal Processing*, 45(2):434–444, 1997.

[18] K. Benchenane, A. Peyrache, M. Khamassi, P. L. Tierney, Y. Gioanni, F. P. Battaglia, and S. I. Wiener. Coherent theta oscillations and reorganization of spike timing in the hippocampal-prefrontal network upon learning. *Neuron*, 66(6):921–36, June 2010.

[19] Y. Benjamini and Y. Hochberg. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society.*, 57(1):289–300, 1995.

[20] H. Berger. Über das Elektrenkephalogramm des Menschen. *Archiv für Psychiatrie und Nervenkrankheiten*, 87(1):527–570, 1929.

[21] 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*, 34:537–541, 1995.

[22] J. Britz, D. Van De Ville, and C. M. Michel. BOLD correlates of EEG topography reveal rapid resting-state network dynamics. *NeuroImage*, 52(4):1162–70, Oct. 2010.

[23] L. Bronge, S.-E. Fernaeus, M. Blomberg, M. Ingelson, L. Lannfelt, B. Isberg, and L.-O. Wahlund. White Matter Lesions in Alzheimer Apolipoprotein E Genotype. *Dementia and Geriatric Cognitive Disorders*, 10:89–96, 1999.

[24] M. J. Brookes, A. M. Gibson, S. D. Hall, P. L. Furlong, G. R. Barnes, A. Hillebrand, K. D. Singh, I. E. Holliday, S. T. Francis, and P. G. Morris. A general linear model for MEG beamformer imaging. *NeuroImage*, 23(3):936–46, Nov. 2004.

[25] M. J. Brookes, J. R. Hale, J. M. Zumer, C. M. Stevenson, S. T. Francis, G. R. Barnes, J. P. Owen, P. G. Morris, and S. S. Nagarajan. Measuring Functional

Connectivity Using MEG: Methodology and Comparison With fcMRI. *NeuroImage*, 56(3):1082–1104, Feb. 2011.

[26] M. J. Brookes, E. B. Liddle, J. R. Hale, M. W. Woolrich, H. Luckhoo, P. F. Liddle, and P. G. Morris. Task induced modulation of neural oscillations in electrophysiological brain networks. *NeuroImage*, 63(4):1918–30, Dec. 2012.

[27] 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):1454–65, Feb. 2007.

[28] M. J. Brookes, J. R. Wood, C. M. Stevenson, J. M. Zumer, T. P. White, P. F. Liddle, and P. G. Morris. Changes in brain network activity during working memory tasks: A magnetoencephalography study. *NeuroImage*, 55(4):1804–1815, Oct. 2011.

[29] M. J. Brookes, M. W. Woolrich, and G. R. Barnes. Measuring functional connectivity in MEG: A multivariate approach insensitive to linear source leakage. *NeuroImage*, 63(2):910–920, Mar. 2012.

[30] M. J. Brookes, M. W. Woolrich, H. Luckhoo, D. Price, J. R. Hale, M. C. Stephenson, G. R. Barnes, S. M. Smith, and P. G. Morris. Investigating the electrophysiological basis of resting state networks using magnetoencephalography. *Proceedings of the National Academy of Sciences of the United States of America*, 108(40):16783–16788, Sept. 2011.

[31] R. Brookmeyer, E. Johnson, K. Ziegler-Graham, and H. M. Arrighi. Forecasting the global burden of Alzheimer's disease. *Alzheimer's & Dementia*, 3(3):186–91, July 2007.

[32] R. L. Buckner, A. Z. Snyder, B. J. Shannon, G. LaRossa, R. Sachs, A. F. Fotenos, Y. I. Sheline, W. E. Klunk, C. A. Mathis, J. C. Morris, and M. A. Mintun. Molecular, structural, and functional characterization of Alzheimer's

disease: evidence for a relationship between default activity, amyloid, and memory. *The Journal of Neuroscience*, 25(34):7709–7717, Aug. 2005.

[33] G. Buzsáki. *Rhythms of the brain*. Oxford University Press, Oxford, 1st edition, 2006.

[34] J. Cabral, E. Hugues, O. Sporns, and G. Deco. Role of local network oscillations in resting-state functional connectivity. *NeuroImage*, 57(1):130–9, July 2011.

[35] R. J. Caselli, E. M. Reiman, D. Osborne, J. G. Hentz, L. C. Baxter, J. L. Hernandez, and G. G. Alexander. Longitudinal changes in cognition and behavior in asymptomatic carriers of the APOE e4 allele. *Neurology*, 62(11):1990–5, June 2004.

[36] A. L. Cohen, D. A. Fair, N. U. F. Dosenbach, F. M. Miezin, D. Dierker, D. C. Van Essen, B. L. Schlaggar, and S. E. Petersen. Defining functional areas in individual human brains using resting functional connectivity MRI. *NeuroImage*, 41(1):45–57, May 2008.

[37] D. Cohen. Magnetoencephalography: Evidence of Magnetic Fields Produced by Alpha-Rhythm Currents. *Science*, 161(3843):784–786, 1968.

[38] D. Cohen. Magnetoencephalography: Detection of the Brain's Electrical Activity with a Superconducting Magnetometer. *Science*, 175(4022):664–666, 1972.

[39] L. L. Colgin. Oscillations and hippocampal-prefrontal synchrony. *Current Opinion in Neurobiology*, 21(3):467–74, June 2011.

[40] P. Comon. Independent component analysis, A new concept? *Signal Processing*, 36(3):287–314, Apr. 1994.

[41] E. H. Corder, A. M. Saunders, W. J. Strittmatter, D. E. Schmechel, P. C. Gaskell, W. Small, A. D. Roses, and J. L. Haines. Gene Dose of Apolipoprotein E Type 4 Allele and the Risk of Alzheimer's Disease in Late Onset Families. *Science*, 261(5123):921–923, 1993.

[42] I. Daubechies, E. Roussos, S. Takerkart, M. Benharrosh, C. Golden, K. D'Ardenne, W. Richter, J. D. Cohen, and J. Haxby. Independent component analysis for brain fMRI does not select for independence. *Proceedings of the National Academy of Sciences of the United States of America*, 106(26):10415–22, June 2009.

[43] F. de Pasquale, S. Della Penna, A. Z. Snyder, C. Lewis, D. Mantini, L. Marzetti, P. Belardinelli, L. Ciancetta, V. Pizzella, G. L. Romani, and M. Corbetta. Temporal dynamics of spontaneous MEG activity in brain networks. *Proceedings of the National Academy of Sciences of the United States of America*, 107(13):6040–5, Mar. 2010.

[44] F. de Pasquale, S. Della Penna, A. Z. Snyder, L. Marzetti, V. Pizzella, G. L. Romani, and M. Corbetta. A cortical core for dynamic integration of functional networks in the resting human brain. *Neuron*, 74(4):753–64, May 2012.

[45] G. Deco and M. Corbetta. The dynamical balance of the brain at rest. *The Neuroscientist*, 17(1):107–123, Feb. 2011.

[46] R. Deichmann. Principles of MRI and Functional MRI. In *fMRI Techniques and Protocols*, chapter 1, pages 3–29. Humana Press, 1st edition, 2009.

[47] R. Duara, D. A. Loewenstein, E. Potter, J. Appel, M. T. Greig, R. Urs, Q. Shen, A. Raj, B. Small, W. Barker, E. Schofield, Y. Wu, and H. Potter. Medial temporal lobe atrophy on MRI scans and the diagnosis of Alzheimer disease. *Neurology*, 71(24):1986–92, Dec. 2008.

[48] M. F. Egan, T. E. Goldberg, B. S. Kolachana, J. H. Callicott, C. M. Mazzanti, R. E. Straub, D. Goldman, and D. R. Weinberger. Effect of COMT Val108/158 Met genotype on frontal lobe function and risk for schizophrenia. *Proceedings of the National Academy of Sciences of the United States of America*, 98(12):6917–22, June 2001.

[49] Elekta Neuromag. Elekta Neuromag System Description, 2006.

[50] Elekta Neuromag. MaxFilter User's Guide. Technical Report October, 2006.

[51] C. Esslinger, H. Walter, P. Kirsch, S. Erk, K. Schnell, C. Arnold, L. Haddad, D. Mier, C. Opitz von Boberfeld, K. Raab, S. H. Witt, M. Rietschel, S. Cichon, and A. Meyer-Lindenberg. Neural mechanisms of a genome-wide supported psychosis variant. *Science*, 324(5927):605, May 2009.

[52] B. S. Everitt and E. T. Bullmore. Mixture model mapping of the brain activation in functional magnetic resonance images. *Human Brain Mapping*, 7(1):1–14, Jan. 1999.

[53] Z. Fatima, M. A. Quraan, N. Kovacevic, and A. R. McIntosh. ICA-based artifact correction improves spatial localization of adaptive spatial filters in MEG. *NeuroImage*, 78:284–294, Apr. 2013.

[54] C. P. Ferri, M. Prince, C. Brayne, H. Brodaty, L. Fratiglioni, K. Hall, K. Hasegawa, H. Hendrie, Y. Huang, A. Jorm, P. R. Menezes, E. Rimmer, M. Scazufca, and Alzheimer Disease International. Global prevalence of dementia: a Delphi consensus study. 366(9503):2112–2117, 2010.

[55] N. Filippini, B. J. MacIntosh, M. G. Hough, G. M. Goodwin, G. B. Frisoni, S. M. Smith, P. M. Matthews, C. F. Beckmann, and C. E. Mackay. Distinct patterns of brain activity in young carriers of the APOE-ϵ4 allele. *Proceedings of the National Academy of Sciences of the United States of America*, 106(17):7209–14, Apr. 2009.

[56] 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):9673–8, July 2005.

[57] K. Friston, L. Harrison, J. Daunizeau, S. Kiebel, C. Phillips, N. Trujillo-Barreto, R. Henson, G. Flandin, and J. Mattout. Multiple sparse priors for the M/EEG inverse problem. *NeuroImage*, 39(3):1104–20, Feb. 2008.

[58] K. Friston, L. Harrison, and W. Penny. Dynamic causal modelling. *NeuroImage*, 19(4):1273–1302, Aug. 2003.

[59] K. J. Friston. Functional and effective connectivity: A review. *Brain*, 2011.

[60] K. J. Friston, K. M. Stephan, J. D. Heather, C. D. Frith, A. A. Ioannides, L. C. Liu, M. D. Rugg, J. Vieth, H. Keber, K. Hunter, and R. S. J. Frackowiak. A multivariate analysis of evoked responses in EEG and MEG data. *NeuroImage*, 3(3 Pt 1):167–74, June 1996.

[61] S. Fujisawa and G. Buzsáki. A 4Hz Oscillation Adaptively Synchronizes Prefrontal, VTA, and Hippocampal Activities. *Neuron*, 72(1):153–65, Oct. 2011.

[62] M. Ganzetti and D. Mantini. Functional connectivity and oscillatory neuronal activity in the resting human brain. *Neuroscience*, 240:297–309, June 2013.

[63] L. U. Gerdes, C. Gerdes, P. S. Hansen, I. C. Klausen, O. Faergeman, and J. Dyerberg. The apolipoprotein E polymorphism in Greenland Inuit in its global perspective. *Human Genetics*, 98(5):546–50, Nov. 1996.

[64] 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):4637–42, Mar. 2004.

[65] E. L. Hall, M. W. Woolrich, C. E. Thomaz, P. G. Morris, and M. J. Brookes. Using variance information in magnetoencephalography measures of functional connectivity. *NeuroImage*, 67:203–12, Feb. 2013.

[66] M. Hämäläinen, R. Hari, and R. J. Ilmoniemi. Magnetoencephalography - theory, instrumentation, and applications to noninvasive studies of the working human brain. *Reviews of Modern Physics*, 65(2):413–497, 1993.

[67] M. Hämäläinen and R. J. Ilmoniemi. Interpreting measured magnetic fields of the brain: minimum norm estimates of current distributions. *Medical and Biological Engineering and Computing*, 32(1):35–42, 1984.

[68] P. Hansen, M. L. Kringleback, and R. Salmelin. *MEG: An Introduction to Methods*. Oxford University Press, Oxford, 1st edition, 2010.

[69] J. Hardy and G. Higgins. Alzheimer's Disease: The Amyloid Cascade Hypothesis. *Science*, 256(5054):184–185, 1992.

[70] B. He, L. Yang, C. Wilke, and H. Yuan. Electrophysiological imaging of brain activity and connectivity - challenges and opportunities. *IEEE Transactions on Biomedical Engineering*, 58(7):1918–31, July 2011.

[71] V. Heise. *How can Magnetoencephalography and Magnetic Resonance Imaging improve our understanding of genetic susceptibility to Alzheimer's disease?* D.phil - under examination, University of Oxford, 2013.

[72] R. N. Henson, E. Mouchlianitis, and K. J. Friston. MEG and EEG data fusion: simultaneous localisation of face-evoked responses. *NeuroImage*, 47(2):581–9, Aug. 2009.

[73] J. Himberg and A. Hyvärinen. ICASSO: Software for investigating the reliability of ICA estimates by clustering and visualization. In *IEEE XIII Workshop on Neural Networks for Signal Processing*, pages 259–268. IEEE, 2003.

[74] J. F. Hipp, D. J. Hawellek, M. Corbetta, M. Siegel, and A. K. Engel. Large-scale cortical correlation structure of spontaneous oscillatory activity. *Nature Neuroscience*, 15:884–890, May 2012.

[75] M. X. Huang, J. C. Mosher, and R. M. Leahy. A sensor-weighted overlapping-sphere head model and exhaustive head model comparison for MEG. *Physics in Medicine and Biology*, 44(2):423–440, 1999.

[76] R. M. Hutchison, T. Womelsdorf, E. A. Allen, P. A. Bandettini, V. D. Calhoun, M. Corbetta, S. D. Penna, J. Duyn, G. Glover, J. Gonzalez-Castillo, D. A. Handwerker, S. Keilholz, V. Kiviniemi, D. A. Leopold, F. de Pasquale, O. Sporns, M. Walter, and C. Chang. Dynamic functional connectivity: Promises, issues, and interpretations. *NeuroImage*, pages 1–20, May 2013.

[77] A. Hyvärinen. Fast and robust fixed-point algorithms for independent component analysis. *IEEE Transactions on Neural Networks*, 10(3):626–34, Jan. 1999.

[78] A. Hyvärinen, J. Karhunen, and E. Oja. *Independent Component Analysis*. Wiley, New York, 1st edition, 2001.

[79] A. Hyvärinen and E. Oja. A Fast Fixed-Point Algorithm for Independent Component Analysis. *Neural Computation*, 9(7):1483–1492, 1997.

[80] R. Ishii, K. Shinosaki, S. Ukai, T. Inouye, T. Ishihara, T. Yoshimine, N. Hirabuki, H. Asada, T. Kihara, S. E. Robinson, and M. Takeda. Medial prefrontal cortex generates frontal midline theta rhythm. *NeuroReport*, 10(4):675–9, Mar. 1999.

[81] C. J. James and C. W. Hesse. Independent component analysis for biomedical signals. *Physiological Measurement*, 26(1):R15–R39, Feb. 2005.

[82] O. Jensen and A. Mazaheri. Shaping Functional Architecture by Oscillatory Alpha Activity: Gating by Inhibition. *Frontiers in Human Neuroscience*, 4:1–8, 2010.

[83] O. Jensen and C. D. Tesche. Frontal theta activity in humans increases with memory load in a working memory task. *European Journal of Neuroscience*, 15(8):1395–1399, 2002.

[84] W. E. Klunk, H. Engler, A. Nordberg, Y. Wang, G. Blomqvist, D. P. Holt, M. Bergstro, I. Savitcheva, M. L. Debnath, J. Barletta, J. C. Price, J. Sandell, B. J. Lopresti, A. Wall, P. Koivisto, G. Antoni, C. A. Mathis, and B. Langstro. Imaging Brain Amyloid in Alzheimer's Disease with Pittsburgh Compound-B. *Annals of Neurology*, 55(3):306–319, 2004.

[85] L. J. Larson-Prior, R. Oostenveld, S. Della Penna, G. Michalareas, F. Prior, A. Babajani-Feremi, J.-M. Schoffelen, L. Marzetti, F. de Pasquale, F. Di Pompeo, J. Stout, M. W. Woolrich, Q. Luo, R. Bucholz, P. Fries, V. Pizzella, G. L.

Romani, M. Corbetta, and A. Z. Snyder. Adding dynamics to the Human Connectome Project with MEG. *NeuroImage*, 80:190–201, Oct. 2013.

[86] Z. Liu, M. Fukunaga, J. A. de Zwart, and J. H. Duyn. Large-scale spontaneous fluctuations and correlations in brain electrical activity observed with magnetoencephalography. *NeuroImage*, 51(1):102–11, May 2010.

[87] H. Luckhoo, M. J. Brookes, V. Heise, C. E. Mackay, K. Ebmeier, P. G. Morris, and M. W. Woolrich. Extracting resting state networks from Elekta Neuromag MEG data using independent component analysis. In *18th Annual Meeting of the Organization for Human Brain Mapping*, 2012.

[88] H. Luckhoo, J. R. Hale, M. G. Stokes, A. C. Nobre, P. G. Morris, M. J. Brookes, and M. W. Woolrich. Inferring task-related networks using independent component analysis in magnetoencephalography. *NeuroImage*, 62:530–541, May 2012.

[89] D. Mantini, R. Franciotti, G. L. Romani, and V. Pizzella. Improving MEG source localizations: an automated method for complete artifact removal based on independent component analysis. *NeuroImage*, 40(1):160–73, Mar. 2008.

[90] D. Mantini, S. D. Penna, L. Marzetti, F. de Pasquale, V. Pizzella, M. Corbetta, and G. L. Romani. A Signal-Processing Pipeline for Magnetoencephalography Resting-State Networks. *Brain Connectivity*, 1(1):49–59, Jan. 2011.

[91] 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):13170–5, Aug. 2007.

[92] L. Marple. Computing the discrete-time "analytic" signal via FFT. *IEEE Transactions on Signal Processing*, 47(9):2600–2603, 1999.

[93] G. Marrelec and J. Doyon. Heading for data-driven measures of effective connectivity in functional MRI. In *Proceedings of International Joint Conference on Neural Networks*, pages 1528–1533, 2005.

[94] G. Marrelec, B. Horwitz, J. Kim, M. Pélégrini-Issac, H. Benali, and J. Doyon. Using partial correlation to enhance structural equation modeling of functional MRI data. *Magnetic Resonance Imaging*, 25(8):1181–9, Oct. 2007.

[95] G. Marrelec, A. Krainik, H. Duffau, M. Pélégrini-Issac, S. Lehéricy, J. Doyon, and H. Benali. Partial correlation for functional brain interactivity investigation in functional MRI. *NeuroImage*, 32(1):228–37, Aug. 2006.

[96] H. Matsuda. Cerebral blood flow and metabolic abnormalities in Alzheimer's disease. *Annals of Nuclear Medicine*, 15(2):85–92, Apr. 2001.

[97] M. J. McKeown, S. Makeig, G. G. Brown, T. P. Jung, S. S. Kindermann, A. J. Bell, and T. J. Sejnowski. Analysis of fMRI data by blind separation into independent spatial components. *Human Brain Mapping*, 6(3):160–88, Jan. 1998.

[98] H. R. Mohseni, M. L. Kringelbach, P. Probert Smith, A. L. Green, C. E. Parsons, K. S. Young, J.-S. Brittain, J. A. Hyam, P. M. Schweder, J. F. Stein, and T. Z. Aziz. Application of a null-beamformer to source localisation in MEG data of deep brain stimulation. In *32nd Annual International Conference of the IEEE Engineering in Medicine and Biology Society.*, pages 4120–3, Jan. 2010.

[99] 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*, 52(4):1149–61, Oct. 2010.

[100] National Institute for Health and Clinical Excellence. Donepezil, galantamine, rivastigmine, and memantine for the treatment of Alzheimer's Disease, 2011.

[101] T. E. Nichols and S. Hayasaka. Controlling the familywise error rate in functional neuroimaging: a comparative review. *Statistical Methods in Medical Research*, 12(5):419–446, Oct. 2003.

[102] T. E. Nichols and A. P. Holmes. Nonparametric permutation tests for functional neuroimaging: a primer with examples. *Human Brain Mapping*, 15(1):1–25, Jan. 2002.

[103] G. Nolte. The magnetic lead field theorem in the quasi-static approximation and its use for magnetoencephalography forward calculation in realistic volume conductors. *Physics in Medicine and Biology*, 48(22):3637–52, Nov. 2003.

[104] P. L. Nunez and S. Ramesh. *Electric Fields of the Brain: The Neurophysics of EEG*. Oxford University Press, Oxford, 2nd edition, 1981.

[105] M. C. O'Donovan, N. Craddock, N. Norton, H. Williams, T. Peirce, V. Moskvina, I. Nikolov, M. Hamshere, L. Carroll, L. Georgieva, S. Dwyer, P. Holmans, J. L. Marchini, C. C. A. Spencer, B. Howie, H. Leung, A. M. Hartmann, H.-J. Möller, D. W. Morris, Y. Shi, G. Y. Feng, P. Hoffmann, P. Propping, C. Vasilescu, W. Maier, M. Rietschel, S. Zammit, J. Schumacher, E. M. Quinn, T. G. Schulze, N. M. Williams, I. Giegling, N. Iwata, M. Ikeda, A. Darvasi, S. Shifman, L. He, J. Duan, A. R. Sanders, D. F. Levinson, P. V. Gejman, S. Cichon, M. M. Nöthen, M. Gill, A. Corvin, D. Rujescu, G. Kirov, M. J. Owen, N. G. Buccola, B. J. Mowry, R. Freedman, F. Amin, D. W. Black, J. M. Silverman, W. F. Byerley, and C. R. Cloninger. Identification of loci associated with schizophrenia by genome-wide association and follow-up. *Nature Genetics*, 40(9):1053–5, Sept. 2008.

[106] G. O'Neill, E. L. Hall, S. Palazzo Corner, P. G. Morris, and M. J. Brookes. A Comparison of Beamformer and Minimum Norm Solutions for Network Connectivity Mapping in MEG. *19th Annual Meeting of the Organization for Human Brain Mapping*, 2013.

[107] J. Onton, A. Delorme, and S. Makeig. Frontal midline EEG dynamics during working memory. *NeuroImage*, 27(2):341–56, Aug. 2005.

[108] T. Ossandon, K. Jerbi, J. R. Vidal, D. J. Bayle, M. A. Henaff, J. Jung, L. Minotti, O. Bertrand, P. Kahane, and J.-P. Lachaux. Transient Suppres-

sion of Broadband Gamma Power in the Default-Mode Network Is Correlated with Task Complexity and Subject Performance. *Journal of Neuroscience*, 31(41):14521–14530, Oct. 2011.

[109] L. Palaniyappan and P. F. Liddle. Does the salience network play a cardinal role in psychosis? An emerging hypothesis of insular dysfunction. *Journal of Psychiatry & Neuroscience*, 37(1):17–27, Jan. 2012.

[110] G. Pfurtscheller, A. Stancak, and C. Neuper. Event-related synchronization (ERS) in the alpha band - an electrophysiological correlate of cortical idling: a review. *International Journal of Psychophysiology*, 24:39–46, Nov. 1996.

[111] M. A. Quraan, S. N. Moses, Y. Hung, T. Mills, and M. J. Taylor. Detection and localization of hippocampal activity using beamformers with MEG: a detailed investigation using simulations and empirical data. *Human Brain Mapping*, 32(5):812–27, May 2011.

[112] 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):676–82, Jan. 2001.

[113] R. Rasetti, F. Sambataro, Q. Chen, J. H. Callicott, V. S. Mattay, and D. R. Weinberger. Altered cortical network dynamics: a potential intermediate phenotype for schizophrenia and association with ZNF804A. *Archives of General Psychiatry*, 68(12):1207–17, Dec. 2011.

[114] E. D. Roberson and L. Mucke. 100 years and counting: prospects for defeating Alzheimer's disease. *Science*, 314(5800):781–4, Nov. 2006.

[115] S. E. Robinson and J. Vrba. Functional neuroimaging by synthetic aperture magnetometry (SAM). In *Recent Advances in Biomagnetism*, pages 302–305, 1998.

[116] Y. B. Saalmann, M. A. Pinsk, L. Wang, X. Li, and S. Kastner. The pulvinar regulates information transmission between cortical areas based on attention demands. *Science*, 337(6095):753–6, Aug. 2012.

[117] J. Sarvas. Basic mathematical and electromagnetic concepts of the biomagnetic inverse problem. *Physics in Medicine and Biology*, 32(1):11–22, Jan. 1987.

[118] J.-M. Schoffelen and J. Gross. Source connectivity analysis with MEG and EEG. *Human Brain Mapping*, 30(6):1857–65, June 2009.

[119] K. Sekihara, S. S. Nagarajan, D. Poeppel, A. Marantz, and Y. Miyashita. Reconstructing spatio-temporal activities of neural sources using an MEG vector beamformer technique. *IEEE Transactions on Biomedical Engineering*, 48(7):760–71, July 2001.

[120] K. D. Singh, G. R. Barnes, and A. Hillebrand. Group imaging of task-related changes in cortical synchronisation using nonparametric permutation testing. *NeuroImage*, 19(4):1589–1601, Aug. 2003.

[121] S. M. Smith, C. F. Beckmann, J. Andersson, E. J. Auerbach, J. Bijsterbosch, G. Douaud, E. Duff, D. A. Feinberg, L. Griffanti, M. P. Harms, M. Kelly, T. Laumann, K. L. Miller, S. Moeller, S. Petersen, J. Power, G. Salimi-Khorshidi, A. Z. Snyder, A. T. Vu, M. W. Woolrich, J. Xu, E. Yacoub, K. Ugurbil, D. C. Van Essen, and M. F. Glasser. Resting-state fMRI in the Human Connectome Project. *NeuroImage*, 80:144–68, Oct. 2013.

[122] 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):13040–5, Aug. 2009.

[123] 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. 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 Suppl 1:S208–19, Jan. 2004.

[124] S. M. Smith, K. L. Miller, G. Salimi-Khorshidi, M. Webster, C. F. Beckmann, T. E. Nichols, J. D. Ramsey, and M. W. Woolrich. Network modelling methods for FMRI. *NeuroImage*, 54(2):875–91, Jan. 2011.

[125] D. Sridharan, D. J. Levitin, and V. Menon. A critical role for the right fronto-insular cortex in switching between central-executive and default-mode networks. *Proceedings of the National Academy of Sciences of the United States of America*, 105(34):12569–74, Aug. 2008.

[126] C. J. Stam and J. C. Reijneveld. Graph theoretical analysis of complex networks in the brain. *Nonlinear Biomedical Physics*, 1(5):3, Jan. 2007.

[127] S. Taulu, M. Kajola, and J. Simola. Suppression of interference and artifacts by the Signal Space Separation Method. *Brain Topography*, 16(4):269–75, Jan. 2004.

[128] C. D. Tesche and J. Karhu. Theta oscillations index human hippocampal activation during a working memory task. *Proceedings of the National Academy of Sciences of the United States of America*, 97(2):919–24, Jan. 2000.

[129] A. J. Trachtenberg, N. Filippini, K. P. Ebmeier, S. M. Smith, F. Karpe, and C. E. Mackay. The effects of APOE on the functional architecture of the resting brain. *NeuroImage*, 59(1):565–72, Jan. 2012.

[130] E. M. Tunbridge, S. M. Farrell, P. J. Harrison, and C. E. Mackay. Catechol-O-methyltransferase (COMT) influences the connectivity of the prefrontal cortex at rest. *NeuroImage*, 68:49–54, Mar. 2013.

[131] N. Tzourio-Mazoyer, B. Landeau, D. Papathanassiou, F. Crivello, O. Etard, N. Delcroix, B. Mazoyer, and M. Joliot. Automated anatomical labeling of

activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *NeuroImage*, 15(1):273–89, Jan. 2002.

[132] P. J. Uhlhaas and W. Singer. Abnormal neural oscillations and synchrony in schizophrenia. *Nature reviews. Neuroscience*, 11(2):100–13, Feb. 2010.

[133] G. Upton and I. Cook. *A Dictionary of Statistics*. Oxford University Press, Oxford, 2nd edition, 2008.

[134] B. D. Van Veen, W. van Drongelen, M. Yuchtman, and A. Suzuki. Localization of brain electrical activity via linearly constrained minimum variance spatial filtering. *IEEE Transactions on Biomedical Engineering*, 44(9):867–80, Sept. 1997.

[135] C. J. Vas, S. Rajkumar, P. Tanyakitpisal, V. Chandra, and World Health Organisation. Azheimer's Disease: The Brain Killer, 2001.

[136] J. Vrba and S. E. Robinson. Signal processing in magnetoencephalography. *Methods*, 25(2):249–71, Oct. 2001.

[137] A. Wimo, L. Jonsson, and B. Winblad. An estimate of the worldwide prevalence and direct costs of dementia in 2003. *Dementia and Geriatric Cognitive Disorders*, 21(3):175–181, Jan. 2006.

[138] W. R. Winter, P. L. Nunez, J. Ding, and R. Srinivasan. Comparison of the effect of volume conduction on EEG coherence with the effect of field spread on MEG coherence. *Statistics In Medicine*, (26):3946–3957, 2007.

[139] M. W. Woolrich, C. F. Beckmann, T. E. Nichols, and S. M. Smith. Statistical analysis of FMRI data. In *fMRI Techniques and Protocols*, chapter 7, pages 179–236. Humana Press, 1st edition, 2009.

[140] M. W. Woolrich, L. Hunt, A. Groves, and G. Barnes. MEG beamforming using Bayesian PCA for adaptive data covariance matrix regularization. *NeuroImage*, 57(4):1466–1479, May 2011.