

Source Modelling of the Human Hippocampus for MEG

Sofie Strunge Meyer

Thesis submitted for PhD

Field of Study: Neuroscience

Supervisor: Gareth Barnes

Submitted October 2016

Awarded December 2016



University College London (UCL)

Declaration:

I, Sofie Strunge Meyer confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Signed:

Date:

Abstract

Magnetoencephalography (MEG) is a neuroimaging technique which gives direct non-invasive measurements of neuronal activity with high temporal resolution. Given its increasing use in cognitive and clinical research, it is important to characterize, and ideally improve upon, its advantages and limitations. For example, it is conventionally assumed to be insensitive to deep structures because of their distance from the sensors. Consequently, knowledge about their signal contribution is limited.

One deep structure of particular interest is the hippocampus which plays a key role in memory and learning, and in organising temporal flow of information across regions. A large body of rodent studies have demonstrated quantifiable oscillatory underpinnings of these functions, now waiting to be addressed in humans. Due to its high temporal resolution, MEG is ideally suited for doing so but faces technical challenges. Firstly, the source-to-sensor distance is large, making it difficult to obtain sufficiently high signal-to-noise ratio (SNR) data. Secondly, most generative models (which describe the relationship between sensors and signal) include only the cortical surface. Thirdly, errors in co-registering data to an anatomical image easily obstruct or blur hippocampal sources.

This thesis tested the hypotheses that a) identification and optimisation of acquisition parameters which improve the SNR, b) inclusion of the hippocampus in the generative model, and c) minimisation of co-registration error, together enable reliable inferences about hippocampal activity from MEG data.

We found the most important empirical factor in detecting hippocampal activity using the extended generative model to be co-registration error; that this can be minimised using flexible head-casts; and that combining anatomical modelling, head-casts, and a spatial memory task, allows hippocampal activity to be reliably observed. Hence the work confirmed the overall hypothesis to be valid. Additionally, simulation results revealed that for a new generation of MEG sensors, ~5-fold sensitivity improvements can be obtained but critically depend on low sensor location errors.

These findings set down a new basis for time-resolved examination of hippocampal function.

This thesis contains work reported in the following manuscripts:

- Paper 1: **Sofie S Meyer**, Holly Rossiter, Matthew Brookes, Mark Woolrich, Sven Bestmann, Gareth R Barnes - Using MEG generative models to make probabilistic statements about hippocampal engagement (under review, NeuroImage)
- Paper 2: **Sofie S Meyer**, James Bonaiuto, Mark Lim, Luzia Troebinger, Holly Rossiter, Sheena Waters, David Bradbury, Simon Little, Sven Bestmann, Matthew Brookes, Gareth R Barnes - Flexible headcasts for high spatial precision MEG (Journal of Neuroscience Methods, 2017)
- Paper 3: **Sofie S Meyer**, Daniel Bush, James A Bisby, Aidan Horner, Neil Burgess and Gareth Barnes - Using head-casts to image hippocampus with MEG (in preparation)
- Paper 4: **Sofie S Meyer**, James Bonaiuto, Daniel Bush, James A Bisby, Aidan Horner, Neil Burgess and Gareth Barnes – MEG sensitivity to hippocampal dipole orientations (in preparation)

Work also undertaken during my PhD but not reported here:

- Elena Boto*, **Sofie S. Meyer***, Vishal Shah, Orang Alem, Svenja Knappe, Peter Krüger, T. Mark Fromhold, Mark Lim, Peter G. Morris, Richard Bowtell, Gareth R. Barnes, Matthew J. Brookes - Room Temperature Magnetoencephalography using Optically-Pumped Magnetometers - (under review, NeuroImage) *denotes equal contribution
- Sheena Waters, **Sofie S Meyer**, Simon Little, James Bonaiuto, Gareth R Barnes and Sven Bestmann - Dissociating lateralised cortical and thalamic sources using high- resolution MEG (in preparation)
- Athina Tzovara, **Sofie S Meyer**, Gareth R Barnes, Raymond J Dolan, Dominik R Bach - Neural oscillations during perception of threat (in preparation)
- James Bonaiuto, **Sofie S Meyer**, Fred Dick, Gareth R Barnes, Sven Bestmann – Laminar specificity of high and low frequency oscillations during action selection - (in preparation)
- Simon Little, James Bonaiuto, **Sofie S Meyer**, Sheena Waters, Gareth R Barnes and Sven Bestmann – Resting state head-cast MEG demonstrates high spatial and temporal resolution (in preparation)

- Holly E Rossiter*, Luzia Troebinger*, James Bonaiuto, **Sofie S Meyer**, Sheena Waters, Simon Little, Sven Bestmann, Gareth R Barnes - Non-invasive evidence from humans that high and low frequency components map to supra- and infra-granular cortex (under revision)
- Elena Boto, Richard Bowtell, Peter Krüger, T. Mark Fromhold, Peter G. Morris, **Sofie S Meyer**, Gareth R. Barnes, and Matthew J. Brookes - The Potential of Optically-Pumped Magnetometers for Magnetoencephalography: A Simulation Study (PLoS ONE, 2016)
- Katja Kornysheva, Daniel Bush, Anna Sadincka, **Sofie S Meyer**, Neil Burgess - Dynamics of spatio-temporal sequence encoding (in preparation)
- Lorenzo Magazzini, Michael Hall, Bethany Routley, Benjamin Hunt, Kevin Prinsloo, **Sofie S Meyer**, Simone Heideman, Rebecca Beresford, Ionannis Papanikolaou, Gareth R Barnes, Matthew Brookes, Paul Furlong, Gary Green, Joachim Gross, Khalid Hamandi, Richard Henson, Vladimir Litvak, Kia Christina Nobre, Mark Woolrich, Krish D Singh - Visual gamma oscillations across the UK: Comparability of UK MEG Partnership data recorded with different MEG scanners (in preparation)

Contents

Abstract	3
List of figures	8
Acronyms and symbols.....	Error! Bookmark not defined.
Acknowledgements	10
Thesis hypotheses and objectives	12
Key contributions	15
Chapter 1 Introduction	16
Origins of the MEG signal.....	18
Technologies for detecting magnetic fields from the brain.....	28
MEG analysis techniques: theory and methods.....	41
Hippocampus and hippocampal oscillations.....	54
Chapter 2 Experiment 1: Using generative models to make probabilistic statements about hippocampal engagement in MEG.....	66
Precis	67
Introduction	67
Hypothesis and objectives.....	69
Methods.....	69
Results.....	76
Discussion.....	92
Key points	96
Chapter 3 Experiment 2: Flexible Headcasts for High Spatial Precision MEG.....	97
Precis	98
Introduction	98
Hypothesis and objectives.....	99
Methods.....	100
Results.....	104

Discussion.....	111
Key points.....	114
Safety procedures.....	114
Chapter 4 Experiment 3: Hippocampal theta activity can be detected in MEG during spatial memory with head-casts	116
Precis.....	117
Introduction	117
Hypothesis	120
Methods.....	120
Results.....	128
Discussion.....	133
Key points.....	135
Chapter 5 Experiment 4: Optimal configuration of optically pumped magnetometers for detecting hippocampal signals	137
Precis.....	138
Introduction	138
Hypothesis and objectives.....	140
Methods.....	140
Results.....	143
Discussion.....	154
Key points.....	155
Chapter 6 Discussion.....	156
General Discussion.....	157
Perspectives and Outlook.....	170
References.....	176

List of figures and tables

Figure 1.1: Human brain, cortical layers, current dipoles and magnetic fields hereof	19
Figure 1.2: Basic structure and electrical response profile of a pyramidal neuron.	23
Figure 1.3: Environmental and biomagnetic noise sources.....	29
Figure 1.4: MEG system set-up and dcSQUID sensor circuitry.	32
Figure 1.5 OPM principles of operation	36
Figure 1.5: Illustration of ill-posed nature of the inverse problem.	43
Figure 1.7 Neural basis for spatial cognition.	58
Figure 1.8 Hippocampal cell morphology and subfield structures.....	59
Figure 2.1: Hippocampal surface structure and location	70
Figure 2.2: Overview of the simulation pipeline.....	73
Figure 2.3: Anatomical models with and without hippocampal priors.	74
Figure 2.4: Sample source reconstructions and model comparison.	78
Table 1: Bayes Omnibus Risk values for hippocampal and cortical simulations	80
Figure 2.5: Anatomical model comparison for hippocampal and cortical (control) sources.	80
Figure 2.6: Effect of co-registration error on anatomical model comparison.....	82
Figure 2.7: Effects of noise and co-registration error on anatomical model comparison.	83
Figure 2.8: Closest cortical neighbour analysis.	85
Table 2: Bayes Omnibus Risk values for hippocampal translations	86
Figure 2.9: Effect of shifting the hippocampal mesh on Free energy.	86
Figure 2.10: Dipole localisation errors as a function of co-registration error and SNR when sources are hippocampal.....	87
Table 3: Sensitivity and Specificity values across co-registration error and SNR levels	89
Figure 2.11: Simultaneous sources.	90
Figure 2.12: Functional model comparison.	91
Figure 3.1: Overview of head-cast construction steps.	102
Figure 3.2 Between-session head movement.	106
Table 4: Mean, maximum and minimum values for coil locations during Experiment 1	107
Figure 3.3 Within-session head movement.	108

Figure 3.4 Consistency of data features across four separate scanning days	111
Supplementary Figure 3A: Head-cast design modified to accommodate hair	115
Figure 4.1: Virtual Reality environment and trial structure.....	122
Figure 4.2: Anatomical models; cortical model, combined model, combined models with shifted hippocampi.....	126
Figure 4.3: Cross validation method	128
Figure 4.4: Free energy differences: cortical versus combined models	129
Figure 4.5: Free energy differences: straight versus rotated hippocampi	130
Table 5: Mean Free energy differences and Bayes Omnibus Risk values across frequency bands/conditions	131
Figure 4.6: Cross validation error: straight versus rotated hippocampi.....	132
Table 6: Cross validation statistics for mesh rotation analysis	132
Figure 5.1 Schematic showing the different types of sensor location error for OPM arrays.	142
Table 7: Bayes Omnibus Risk values for OPM and SQUID simulations across source strengths	143
Figure 5.2 ΔF values across simulated source strengths for OPM and SQUID arrays.....	144
Figure 5.3 Effects of systematic and independent sensor location errors on ΔF across simulation strengths.....	146
Figure 5.4 Evidence for source models as an effect of sensor location error	148
Figure 5.5: Effects of systematic and independent sensor orientation errors on ΔF across simulation strengths.....	149
Figure 5.6 Evidence for source models as an effect of sensor orientation error	151
Figure 5.7 Evidence for combined generative model as sensors degrade	153
Figure 6.1 OPM sensor placement for optimal detection of hippocampus.....	173

Acronyms and symbols

BEM	Boundary Element Method
BOR	Bayes Omnibus Risk
CTF	Canadian Thin Films
DLE	Dipole Localisation Error
EBB	Empirical Bayes Beamformer
EEG	Electroencephalography
F	Negative variational Free energy*
fMRI	Functional Magnetic Resonance Imaging
fT	femto Tesla
GS	Greedy Search
h	Vector of hyperparameters
I	Current
J	Current density
L	Lead field matrix
MEG	Magnetoencephalography
MNE	Minimum Norm Estimate
MRI	Magnetic Resonance Imaging
MSP	Multiple Sparse Priors
OPM	Optically Pumped Magnetometry/Magnetometer
PEB	Parametric Empirical Bayes
PET	Positron Emission Tomography
PSPs	Post synaptic potentials
Q	Prior source covariance matrix
ReML	Restricted Maximum Likelihood
RF	Radio frequency
SERF	Spin Exchange Relaxation Free
sLORETA	Standardized low resolution brain electromagnetic tomography
SNR	Signal-to-Noise Ratio
SQUIDs	Superconducting Quantum Interference Devices
Y	MEG data

*Throughout this thesis, ‘model evidence’, and ‘Free energy’ will be used interchangeably. Free energy is an approximation of model evidence.

Acknowledgements

First, I would like to thank my primary supervisor, Gareth Barnes. Thank you for countless hours of mentoring, teaching and fruitful discussion. Thank you for all the snippets of code, and for always leaving me feeling more motivated and excited after meetings. Thank you for letting me give lots chances to give talks and attend conferences and workshops, collaborate with other great people, and for the nuggets of wisdom you've shared. Thank you for being so patient while I have been learning how to write, scan and think. And perhaps most symbolically, thank you for all the high-fives you've given me in return for new plots!

This thesis would not have been possible without the help and support of many good friends and colleagues from UCL who have generously given me time, good advice, help and code. Thank you to Luzia Troebinger, Sven Bestmann, James Bonaiuto, Jose Lopez, Ryszard Aukstulewicz and Saskia Helbling for help and hints on data analysis, meshes, math and life. Thank you to Yoshihito Shigihara, Letty Manyande, David Bradbury and Al Reid for help scanning. Thank you to the great IT team at the FIL for always being helpful, patient and thorough. Thank you to Desmond Lim and Mark Lim for invaluable creativity and problem-solving for designing and 3D printing head-casts in many shapes and sizes. Thank you to Dan Bush (also for helpful feedback), Aidan Horner, James Bisby, and my second supervisor Neil Burgess for help designing an experiment that worked and looks beautiful. Thank you to my dear office mates who have kept me (somewhat) sane and without whom I wouldn't have had nearly as much fun – thank you in particular to Tobi Hausser, Andrea Gajardo Vidal, Iris Vilares and Biswa Sengupta. Thank you to the FIL for being a great place to work and for providing me with free strong coffee and wifi for three years. Thank you to Krish Singh for accepting me onto the partnership programme. Thank you to the Medical Research Council and Engineering and Physical Sciences Research Council (grant MR/K501086/1).

Thank you to my beloved family, Mom, Dad, Josi, Olli, for all the support, encouragement and proofreads. Tak!

And thank you to Zeb Kurth-Nelson for way too many things to list on one page.

Thesis hypotheses and objectives

The over-arching hypothesis addressed in this thesis is that it is possible to improve the detection of hippocampal activity in MEG data. This hypothesis is two-fold: 1) It is possible to improve data acquisition by improving the signal-to-noise ratio (SNR) through the use of head-casts to stabilize the relationship between MEG sensors and the brain. 2) It is possible to improve data analysis by including more anatomical and electrophysiological information about the hippocampus.

The general aim of this thesis has therefore been to construct, critically evaluate, and apply hippocampal source models to MEG data. Adding anatomical detail to the generative model allows us to explicitly test how well this can help us explain variance in MEG data. Further, it enables us to directly quantify when and how empirical obstacles such as poor SNR and co-registration error impede this ability (Experiment 1). Knowledge of these obstacles nevertheless leaves open questions. One of the starkest is perhaps “but does it work?” To address this, we therefore empirically validate the new generative model using a task that has been repeatedly demonstrated to engage hippocampus (Experiment 3).

In parallel, another aim is to increase the SNR during data acquisition to get less noisy and more spatially accurate, precise, and reproducible signals from the hippocampus. First, we develop flexible head-casts to be used in combination with conventional MEG recording (Experiment 2), and later, through data-based optimization of the sensor configuration for hippocampal activity, we optimise the SNR with a new generation of room-temperature MEG sensors (Experiment 4).

Experiment 1: Probabilistic statements about simulated hippocampal activity using generative models

Hypothesis: If the hippocampus is explicitly incorporated into the generative MEG source model, then it is possible to test whether or not it is active at a certain time and within a certain frequency band. The validity of this hypothesis can be tested using simulated data (where the ground truth is known), making it possible to identify the extent to which different empirical acquisition factors - such as co-registration error and SNR - hinder the detection of hippocampal activity.

Experiment 2: Flexible head-cast design for high spatial precision MEG

Hypothesis: If the co-registration error and head movement can be minimized during acquisition, then the SNR will be significantly improved, leading to much better quality data with more consistency and less variability both across and within recording sessions. We also predict that such improvement of the SNR will improve the reproducibility of the data and spatial resolution of the inference in general.

Experiment 3: Empirical MEG recordings of hippocampal activity using head-casts and hippocampal source models

Hypotheses: 1) If we combine an acquisition technique which is optimal for obtaining high SNR data (Experiment 2, head-casts), a well-validated spatial memory task which is known to engage the hippocampus, and explicit source modelling of the hippocampus (Experiment 1), we can detect hippocampal sources in real MEG data.

2) If this combination of tools is effective, then changes to the hippocampal portion of the generative model should give rise to decreases in model generalizability/fitness (which can be quantified by two orthogonal metrics; Free energy and cross-validation error). Specifically, we predict that if the subject-specific generative model of the hippocampus is correct, then laterally rotating it should decrease the model evidence and increase the cross validation error

Experiment 4: Optimization of acquisition parameters to detect hippocampal activity using Optically Pumped Magnetometers

Hypothesis: If we can utilize a new generation of room temperature (as opposed to supercooled) MEG sensors to drastically improve the SNR due to decreased source-to-sensor distances, then we can in turn optimize the configuration (location and orientation) of sensors based on results obtained in the first three experiments to optimally detect hippocampal signals. Such detection of hippocampal signals will allow the possibility of eventually making MEG recordings of the hippocampus as the participant moves freely around the environment.

In summary, the main aims of this thesis are a) modelling of the hippocampus as a potential electromagnetic source giving rise to part of the MEG signal, b) characterization of the empirical requirements for detecting the signal originating from hippocampus, and c) optimization of the acquisition parameters in order to meet these requirements.

Modelling the signal consists of extending the existing biophysical model used to explain MEG data, by including the location, orientation, density, cell type and global geometrical shape of the hippocampus.

By systematically simulating a set of MEG experimental data, the aim is to first characterize the requirements for hippocampal signal detection. This is then carried forward to designing and conducting a cognitive experiment and head-cast device which meets the requirements identified in simulations. Finally, we explore where on the surface of the head these signals are strongest and use this to guide the placement of room-temperature sensors.

Key contributions

- The work provides a model of the hippocampus as an electromagnetic signal generator and several novel ways of testing this model. We found that including a hippocampal ‘mesh’ (i.e. anatomical surface) within the standard model of the brain helps explain hippocampal source activity but does not introduce bias. We applied rigorous control tests to determine the spatial specificity and limitations and advantages of including this mesh.
- Development of a new prototype of flexible head-casts, which minimizes both head movement during scanning, and errors in co-registration to anatomical data. Importantly, these casts are safer and more comfortable than the previous design. Unlike the previous head-casts, this prototype also enables subjects to see while being scanned.
- Demonstration of hippocampus-specific activity recorded with MEG. Data was acquired by asking subjects to perform a cognitive task known to evoke hippocampal theta band oscillations while they were wearing a flexible head-cast. Through application of Bayesian model comparison and cross-validation, we found that lateral rotations of the hippocampal portion of the generative model significantly decreased its predictive power, even when these errors were as low as 5°.
- The PhD work also contributed novel conceptual and theoretical ideas for efficient use of room-temperature MEG sensors. These sensors represent a new potential for MEG research to have higher spatial resolution through a 5-10-fold SNR improvement. The empirical requirements for detecting hippocampal activity in terms of different sources of error are addressed and the spatial topography of a hippocampal source was obtained, giving way for constructing the spatial configuration of a hippocampus-optimal sensor array.

Chapter 1

Introduction

Overview

Deep brain structures such as the hippocampus are involved in many healthy and pathological brain processes in humans, yet are relatively poorly understood in the context of temporal dynamics. Despite the potential and although clinical, simulation-based, and empirical demonstrations of the detectability of hippocampus using magnetoencephalography (MEG) have been presented (Attal and Schwartz, 2013; Korczyn et al., 2013; Papanicolaou et al., 2005), widespread scepticism against the possibility for detecting hippocampal activity by MEG persists (Mikuni et al., 1997; Riggs et al., 2009; Stephen et al., 2005). This thesis aims to establish, from a methodological standpoint, the detectability of hippocampal signals in MEG. MEG detection of hippocampal signals opens up exciting possibilities to formulate and test new and specific hypotheses about the dynamics of hippocampal function during cognitive functions in which it is known to be engaged, such as episodic memory and spatial navigation.

MEG is a non-invasive neuroimaging technique that measures electromagnetic brain activity with millisecond temporal resolution. In order to localise the spatial origin of such activity, anatomical and electrophysiological information is used to constrain the solution space. Whilst this general framework is well-established for neocortical sources (Hämäläinen et al., 1993; Henson et al., 2009; Lopes da Silva, 2013; Vrba and Robinson, 2001), reconstruction of deep sources remains controversial (Hämäläinen et al., 1993; Mikuni et al., 1997; Riggs et al., 2009; Stephen et al., 2005). This is partly because the signal strength - and consequently the spatial resolution - rapidly decreases with distance from sensors: $\text{strength} \propto 1/\text{distance}^2$ for dipoles (Geselowitz, 1967; Hämäläinen et al., 1993; Hillebrand and Barnes, 2002; Sarvas, 1987), and partly because it is unclear whether particular cell features or configurations of deeper structures render them magnetically silent (Hämäläinen et al., 1993). Despite a well-characterized repertoire of characteristic oscillatory dynamics (for reviews see Buzsáki, 2006; O'Keefe, 2007), it is often assumed that the hippocampus is difficult to detect with MEG, an assumption that has only recently begun to receive critical reappraisal (Attal and Schwartz, 2013; Riggs et al., 2009).

Research on the methods required to characterize hippocampal detectability has not been matched by attention to fMRI, rodent, behavioural, and intracranial methods used to characterize hippocampal functions. Here, we are concerned with the analysis and acquisition parameters required for successful and robust non-invasive detection

of human hippocampal signals. The general relevance of this aim extends to clinical, computational, and cognitive research questions.

In this thesis this issue is addressed by employing a range of different techniques including simulations, anatomical and electrophysiological modelling, acquisition optimisation through design, virtual reality environments, and room temperature optically pumped magnetometers.

In this introductory chapter, an overview of MEG as a neuroimaging modality is presented, starting with the origins of the MEG signal, and how this signal can be detected with the two different types of sensors this project used. Next the theory and mathematical methods used to characterise the observed signal in 3D space are introduced. Generally, these methods rely on specific assumptions about the magnetic signal: both where it might originate, and how it might co-vary between neighbouring locations. To tie these concepts together with the methods employed in later chapters, the introduction describes how such assumptions can be formulated as hypotheses, and how these can be directly compared in a Bayesian framework. Finally, the relevance of, and rationale for, studying the hippocampus using MEG is outlined, and both the modelling and simulation approaches used are described. In addition, the debate regarding detectability of hippocampus using MEG is briefly reviewed, and several previously unresolved questions that are addressed in this thesis are highlighted.

Origins of the MEG signal

Electromagnetism of the brain

The brain transmits information through electrical activity and electrical current flow gives rise to magnetic fields. MEG measures these magnetic fields. The MEG signal is thought to originate mainly in the outermost layer of the brain - the cerebral cortex - which consists of a 2-4 mm thick sheet of grey matter. The surface area of this sheet in spread out form is $\sim 2500 \text{ cm}^2$ for an adult human brain (Hämäläinen et al., 1993). In order to fit it inside the skull, it is therefore highly folded (**Figure 1.1A**). Interestingly, the cellular architecture is well-preserved across this sheet: it is possible to subdivide it into different layers based on the morphological features of cells within it (**Figure**

1.1B). Although the absolute thickness of the sheet and the layers vary to an extent with function of a particular brain area, the layers are generally well-preserved. Of main interest here are the layers containing a particular type of neuronal cell, namely pyramidal cells. These cells are large, recurrently connected (i.e. excite and inhibit one another so as to potentially introduce synchronous activation), and oriented in parallel. Pyramidal cells are found mainly in layers II/III and V, with those in layer V being larger and longer (**Figure 1.1**). Layer V is therefore thought to constitute the main contribution to the recorded MEG signal (Murakami and Okada, 2006) (**Figure 1.1**).

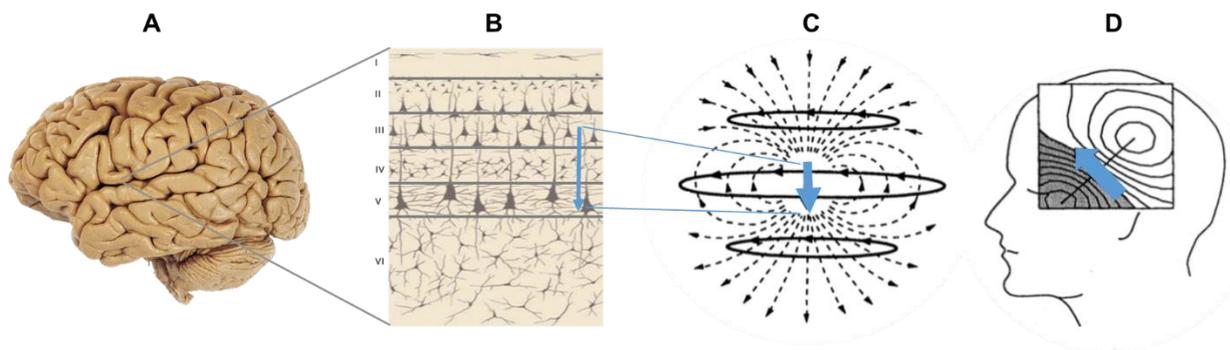


Figure 1.1: Human brain, cortical layers, current dipoles and magnetic fields hereof.

A) Lateral view of the human brain. The cortical surface is intricately folded, allowing more surface area and therefore computational power inside the skull. Photograph adapted from [Visuals Unlimited]. **B)** Layering of different cell types in outermost sheet of cortex (cerebral cortex). Horizontal lines mark subdivision into six distinct layers. These are, from outer to innermost layers the molecular (I), external granular (II), external pyramidal (III), internal granular (IV), internal pyramidal (V) and multiform (VI). Together these six layers constitute the grey matter. Below them are white matter and above them are the pia, arachnoid, and dura matter before the skull. Layers II/III and V contain pyramidal cells which have triangularly shaped cell bodies as shown. Blue arrow represents a current dipole produced by electrical activity in a layer V pyramidal cell. Diagram adapted from <https://o.quizlet.com/X.dubvVJnupqdHtsbEOi9A.png> **C)** Magnetic field of a current dipole. Blue arrow represents current dipole created by a primary current, dotted lines represent the volume current balancing it and solid lines represent the magnetic field. Note that the magnetic field is technically continuous and the width of the circles represent its strength at three heights. The magnetic fields “wrap around” the current dipole and should therefore be interpreted as a three-dimensional ring around the blue arrow. This diagram assumes the conducting medium to be homogeneous which is largely true for magnetic fields in the brain. **D)** Topographical field map derived from MEG signals. The maxima and minima of the magnetic fields represent the locations where the strongest part of the magnetic field exits and enters the skull. The dipole is midway between these two points. Images C and D adapted from (Hämäläinen et al., 1993).

In the next section the production of a *primary* current (blue arrow, **Figure 1.1B,C**) is described in terms of movement of ions and thus charge along the membrane of pyramidal cells. To compensate for this and prevent build-up of charge, a passive *volume* current (dotted lines, **Figure 1.1C**) restores the balance of charge through ion flow in the opposite direction in the extracellular space. Both currents contribute to the resulting magnetic field (solid lines, **Figure 1.1C**) and can be approximated as a current dipole: two electrical charges separated by a small distance (such as the length of a neuron), with equal magnitude but opposite charge (blue arrow, **Figure 1.1B-D**). Here we focus first on the physics and neurobiology of pyramidal cells and current dipoles, and how these give rise to the measured MEG signal.

Electrical and chemical signalling of neurons

The nervous system consists of billions of specialized cells which have evolved to carry and transfer information. While the former is achieved electrically within nerve cells or neurons, the latter is achieved chemically when passing information from one neuron to the next. Both of these processes are mediated by the opening and closing of ion channels in the cellular membrane.

Within-neuron communication relies on well-maintained electrochemical gradients across the membrane which, when altered, result in fluctuations in the local membrane potential. The main ions involved are sodium (Na^+), chloride (Cl^-), calcium (Ca^{2+}), and potassium (K^+). While Na^+ , Cl^- and Ca^{2+} have higher extracellular concentrations, K^+ has a higher intracellular concentration. These chemical gradients are balanced by electrical gradients. For each ion, there exists a membrane potential at which the two gradients are exactly balanced and there is no net flow. This is known as the reversal potential. It is the value of the reversal potential relative to the neuron's 'resting' potential (around -65 mV) that determines whether an increase in ionic permeability due to channel opening will result in de- or hyper- polarisation of the membrane. When the value of the membrane potential is lower than resting potential (i.e. -70 and -110 mV for Cl^- and K^+ respectively), the membrane will hyperpolarise if the channels open. Conversely, when it is higher (i.e. +40 and 0 mV for Na^+ and Ca^{2+} respectively), the membrane will depolarise when the channels open. When the channel is depolarised past a certain threshold (about -55 mV), the change in electrical potential becomes an absolute and highly stereotyped potential which propagates along the length of the

membrane and is referred to as an action potential. This propagation of the action potential continues along the axon until it reaches the synapse, where it causes release of neurotransmitters at the axonal terminals.

Between-neuron communication relies on local release of neurotransmitters into the synaptic cleft between cells. These neurotransmitters act on, and thereby open, ligand-gated ion channels on the neighbouring neuron. Similar to within-cell communication, small electrical potentials are generated as a result of ion flux when the channels open. These potentials are commonly referred to as post-synaptic potentials (PSPs) and form the basis of the signal detected with MEG. As neurons can either excite or inhibit one another (or themselves), two main classes of neurotransmitters exist: excitatory and inhibitory. The main excitatory pathway involves the neurotransmitter glutamate which acts on either AMPA¹ (Na⁺ permeable), or NMDA (Ca²⁺/Na⁺ permeable) receptors. This gives rise to excitatory (i.e. depolarising) post synaptic potentials (EPSPs). Conversely, inhibition most commonly works through the release of the neurotransmitter GABA which acts on so-called GABA receptors (Cl⁻ permeable) and generates inhibitory (i.e. hyperpolarising) post synaptic potentials (IPSPs). Because these potentials are a result of inputs from other cells, they occur mainly on the dendrites of the neuron; but importantly, the dipole used to model these PSPs spans the length of the neuron, as the current sources (outward currents) and sinks (inward currents) are located at opposite ends of the neuron (**Figure 1.2A**).

Four important differences between action potentials and PSPs distinguish them in terms of detectability in MEG. First, action potentials are biphasic whereas EPSPs are monophasic. Therefore, when detecting activity synchronized across large cell populations (around 10⁴), action potentials may cancel each other out if not exactly synchronized, whereas EPSPs summate as long as they overlap in time. Second, EPSPs lend themselves well to detection in MEG because they are slower and thus have a larger window during which these overlaps can take place. While action potentials typically last only around a millisecond, EPSPs last tens of milliseconds, depending on the receptor type – the decay time is ~2 ms for AMPA receptors and up to 100 ms for NMDA receptors (Spruston et al., 1995) (**Figure 1.2B**). Third, an action potential consists of de- and re-polarisation moving along the length of the axon. This

¹ AMPA stands for α -amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid, NMDA stands for N-methyl-D-aspartate, and GABA stands for γ -aminobutyric acid.

is modelled well by two opposing dipoles; one leading and one trailing the depolarisation. If the conduction velocity of an axon is 1 m/s, then the distance between two such dipoles will be ~ 1 mm (Hämäläinen et al., 1993). Consequently, they will form a quadrupole at a distance (Plonsey, 1977), which means that the contribution will be far weaker than dipolar sources such as EPSP-induced neuron-wide potentials. Specifically, the decay of the magnetic field strength with distance from the primary current follows the inverse square law. Therefore, for distance r_d between a given current source and a sensor detecting its field at a point in space, the decay is proportional to $1/r_d^2$ if the current is a dipole, and $1/r_d^3$ if it is a quadrupole (Hämäläinen et al., 1993).

Because the magnetic field associated with action potentials is quadrupolar, it will decay more rapidly with distance and therefore be less detectable than dipolar elements at a distance – such as the scalp. Finally, the apical dendritic trees (**Figure 1.2A**) of neighbouring pyramidal cells tend to be aligned and thus lend themselves well to formation of dipolar activations across a population of many neurons, whereas the same is not true for the axons in such a population (except when they enter white matter pathways). Generation of measureable fields also depends on recurrent connectivity across the neuronal circuits – i.e. pyramidal cells are structurally configured to do so but also require synchronisation, e.g. through recurrent connectivity between them, or with the help of interneurons which help synchronisation. Thus, functional and structural connections are needed for synchronising populations. Such circuits can be found in the pyramidal cell layers of both neocortex (layers II/III and V) and the hippocampus. In the hippocampus, a single pyramidal cell layer which is morphologically highly similar to neocortical layer V spans the Ammon's horn subfields (Amaral and Witter, 1989) (see also *Hippocampus and Hippocampal Oscillations* section). Thus, the 'open field' arrangement underpinning the generation of macroscopic electrical potentials can be found across dendrites, but not in axons along which the dipoles cancel out. Critically, these parallel dendritic arrangements are also perpendicular to the surface of the cortical sheet, which means that when it is tangential to the surface of the head, the magnetic field is detectable outside the head (although radial dipoles are therefore lost, these make up a very small proportion of the potential cell assemblies (Hillebrand and Barnes, 2002)).

In summary, synchronous excitatory input to apical dendrites of a population of parallel pyramidal cells induces EPSPs that can be modelled as a dipole moment spanning the height of the neuron cluster, as the sinks and sources are maximally separated. Given that these magnetic fields are dipolar, they are detectable at a distance.

Furthermore, they are likely to summate sufficiently due to their monophasic nature and relatively slow time constants, as well as parallel structural arrangements (**Figure 1.2C**).

There are nonetheless instances where action potentials may contribute to the measured signal, such as during epileptic seizures (Bragin et al., 2002), at very high frequency activity (>100 Hz) (Curio, 2000), or from somatosensory stimulation (in rat neocortex) (Barth, 2003). However, the focus of this thesis is on modelling the signals observed at much lower (e.g. 4-8 Hz) frequencies, and it is highly unlikely that action potentials influence these signals.

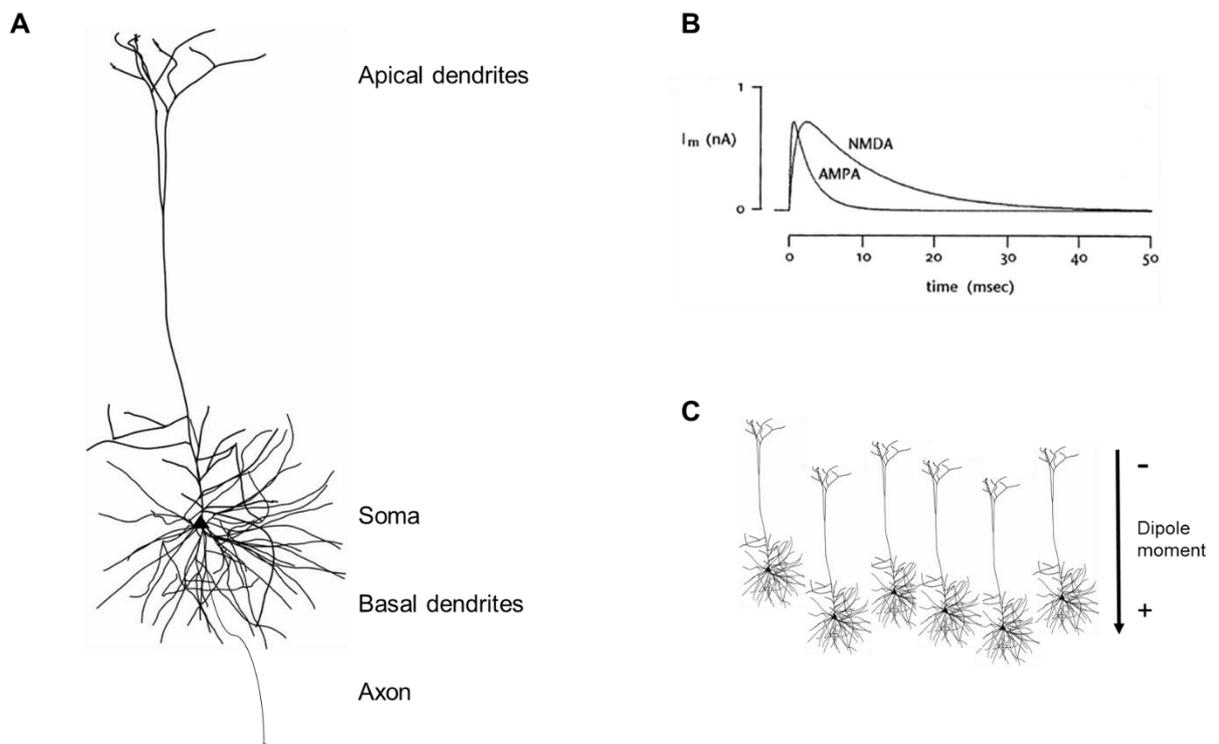


Figure 1.2: Basic structure and electrical response profile of a pyramidal neuron.

A) The neuron receives inputs from other neurons on its dendritic tree, on both the basal and apical portions. Excitatory inputs cause excitatory postsynaptic potentials (EPSPs) to arise. Reconstructed morphology of layer 5 pyramidal cell, adapted from <http://bluebrain.epfl.ch/> **B)** EPSPs vary in duration depending on the receptor type. I_m represents the membrane current, quantified in nano Amperes (nA), following a unitary synaptic input at $t=0$ ms. From (Shepherd, 2003) **C)** Example of an open field: pyramidal neurons arranged in parallel with a dipole moment arrow next to it. The neurons receive synchronous excitatory inputs to the apical dendrites. The sink is therefore at the top or near the apical dendrites, while the source is at the bottom near the soma.

A subtle but important point is the relationship between distributed excitatory synaptic inputs to dendrites (Spruston, 2008), and the modelling of this activity as an equivalent current dipole (ECD). There are complex and specific cytoarchitectonic differences between the synaptic targets of different sub- and neo-cortical projections, which mean that the dendritic arbors contact a given pyramidal cell at different points. Consequently, the distance between the sink and the source of a given dipole will differ depending on where the excitatory input was delivered, causing differences in the magnetic field that is generated. For example, excitation of the soma of layer III pyramidal cells gives a positive surface potential, whereas excitation of the apical dendrites of layer V pyramidal cells produces a negative surface potential (Mitzdorf, 1985). Hence, it is difficult to draw conclusions about whether the underlying signal is excitatory or inhibitory in nature based on the sign of the cortical dipole.

Electromagnetic coupling and properties of the signal

As different patches on the cell membrane act as current sources (outward currents) and sinks (inward currents), magnetic fields are generated around the current flow. The EPSPs of neurons thus underlie the measured signal, but these are not strong enough to be individually detected at the scalp. One estimate is that the current dipole of a 2 mm long cortical pyramidal cell is between 20-200 fAm (Hämäläinen et al., 1993; Murakami and Okada, 2006), while the weakest signal measurable is around 10 nAm (Hämäläinen et al., 1993). It follows that if one measures a 20 nAm dipole, the number of concurrently active cells generating the signal must be between 100,000 and 1,000,000.

The Danish physicist Hans Christian Ørsted was the first to describe the link between electric currents and magnetic fields in 1819-20. Before this, in 1786, Luigi Galvani demonstrated the presence of electrical current in animal (specifically frog) tissue. Later, in 1831, James Clerk Maxwell proposed a system of partial differential equations describing how electrical (E) and magnetic fields (B) are generated by the rate of change of each other, and the presence of charge density (ρ) and current density (J). He also proposed that the propagation of the electrical and magnetic fields could be described with a single wave equation, and that the speed of this propagation is equal to that of light. Since MEG sensors are only a few centimetres (or millimetres in the case of newer sensors) from the brain, the delay between generation and sensor

detection of electromagnetic activity need not be considered. Notably, this is different from other neuroimaging modalities such as positron emission tomography (PET) or functional Magnetic Resonance Imaging (fMRI) where the signal can only be detected after several minutes or seconds, respectively. The second reason why Maxwell's equations are useful for describing MEG data, is that brain currents have sufficiently low frequencies (<1000 Hz), such that the capacitance and impedance of the head and brain tissues, the inductive effect (dipoles across chemical bonds) and the electromagnetic propagation effects (attenuation, reflection or perturbations of the waves) are all negligible. Together, the instantaneous propagation and the relatively low frequencies mean that the time derivatives in the original equations can be ignored and thus the equations take a simpler, 'quasi-static' form (Geselowitz, 1967; Hämäläinen et al., 1993).

The key notion of electromagnetic coupling as described by Ørsted, is that an electric current produces a circular magnetic field as it flows, be it through a wire, or for the purposes of understanding MEG, through a neuron. This primary current generates magnetic fields that are transmitted through the biological tissue of the head and towards sensors outside. Critically, this relationship is not only quasi-static, but also linear such that the weighted sum of a subset of currents is equal to the weighted sum of their corresponding magnetic fields (Kutas and Dale, 1997). The Biot-Sarvas law describes this relationship between a current density $J(r')$ at location r' in relation to its magnetic field $B(r)$ at a given location r (Baillet et al., 2001; Hallez et al., 2007)

$$B(r) = \frac{\mu_0}{4\pi} \int J(r') \times \frac{r - r'}{\|r - r'\|^3} dv' \quad (1.1)$$

Where μ_0 is the permittivity of free space (a fundamental constant, (Baillet et al., 2001; Hämäläinen et al., 1993)), and dv' is the differential element of volume. The current $J(r')$ depends both on primary current flow (in/along the neuronal membrane), and on volume current flow (compensatory ion flow in the extracellular space). In MEG, the more relevant of the two is the primary current, as the location hereof is the location of the active neuron assembly. For a given distribution of primary currents and potentials on all surfaces of interest (i.e. patches of cortex with sufficiently large cell assemblies), the magnetic field can be directly calculated, and becomes (Geselowitz, 1967):

$$B(r) = B_0(r) + \frac{\mu_0}{4\pi} \sum_{ij} (\sigma_i - \sigma_j) \int_{S_{ij}}^k V(r') \frac{r - r'}{\|r - r'\|^3} \times dS_{ij} \quad (1.2)$$

Where $B_0(r)$ is the magnetic field due to primary currents alone, summed over all boundaries (inner skull surface, outer skull surface, scalp). The second term describes the contribution of the volume current to the magnetic field, in the form of surface integrals across the brain-skull, skull-scalp, and scalp-air boundaries. σ denotes the conductivity term, which is assumed to be isotropic and constant for each of the three surfaces, and $V(r')$ denotes the potential at r due to the primary current. S_{ij} are the modelled surface finite elements.

This equation thereby states that the magnetic field can be calculated directly, given the primary current distribution and the potential on all surfaces. In other words, it is hereby possible to compute the MEG signal generated by would-be neural activity. This is also known as forward modelling or solving the forward problem. In turn, this enables inferences about the spatial origins of the recorded MEG signal. To compute the forward model, we need to specify a primary current distribution $J^p(r')$ from which we can calculate the primary magnetic field B_0 :

$$B_0(r) = \frac{\mu_0}{4\pi} \int J^p(r') \cdot \frac{r - r'}{\|r - r'\|^3} dr' \quad (1.3)$$

The forward problem is solved (or the forward model is provided) by using the primary magnetic field $B_0(r)$ to model the external magnetic fields. The next two sections deal with how the magnetic fields are detected and how the primary current distribution can be modelled (forward modelling).

Summary

The signal measured in MEG primarily originates from postsynaptic potentials (PSPs) in pyramidal cells. Synchronized excitatory inputs summate and give rise to dipolar magnetic fields which are detectable at the scalp. Important structural features are parallel and scalp-tangential arrangements of the dendritic trees, which is the case in the pyramidal cell layers found in layers II/III and V of neocortex, as well as the single pyramidal cell layer in hippocampus.

The most important feature of the MEG signal is that, because of the instantaneous propagation of magnetic fields from the brain to the sensors, it presents a direct but non-invasive measure of real-time neural activity.

Technologies for detecting magnetic fields from the brain

In order to successfully record the very small magnetic signals generated by the brain, one must both reduce the otherwise overpowering external magnetic fields, and use highly sensitive sensors. Further, it is appropriate to configure the sensors to have a high sampling rate (many Hz) because the signals reflect real-time brain activity, unlike with for example fMRI. This section describes different currently used noise reduction techniques and the two different types of magnetometer sensors used in this thesis. While the conventional Superconducting Quantum Interference Device (SQUID) based sensors require low temperature environments to function, the newer generation of Optically Pumped Magnetometer (OPM) sensors require heating but only in a small, sub-cubic centimetre, sensitive volume. In practice this means that they can be approximately room temperature a few millimetres from this volume and therefore from the scalp. This section briefly describes the quantum mechanics and physics underpinning these sensors.

Noise reduction

The magnetic fields generated by the brain are extremely weak. As mentioned, thousands or hundreds of thousands of cells are required to give rise to a measureable signal. Nonetheless, the amplitude of a typical evoked (stimulus time-locked) response is only ~100 femtoTesla (fT, 10^{-15} T). This is around one billion times smaller than the constant, ~50, μ T magnetic field of the earth (**Figure 1.3**). Additionally, other large signals come from cars, trains, computers, people, power lines, and metallic doors – all of which are typically found within short distances of an MEG scanner. Finally, magnetic fields generated by the heart or eyes can in some cases overpower the brain signal, even if this is detected at the scalp. It follows that in order to detect signals from the brain, the external signals must be reduced or ideally eliminated.

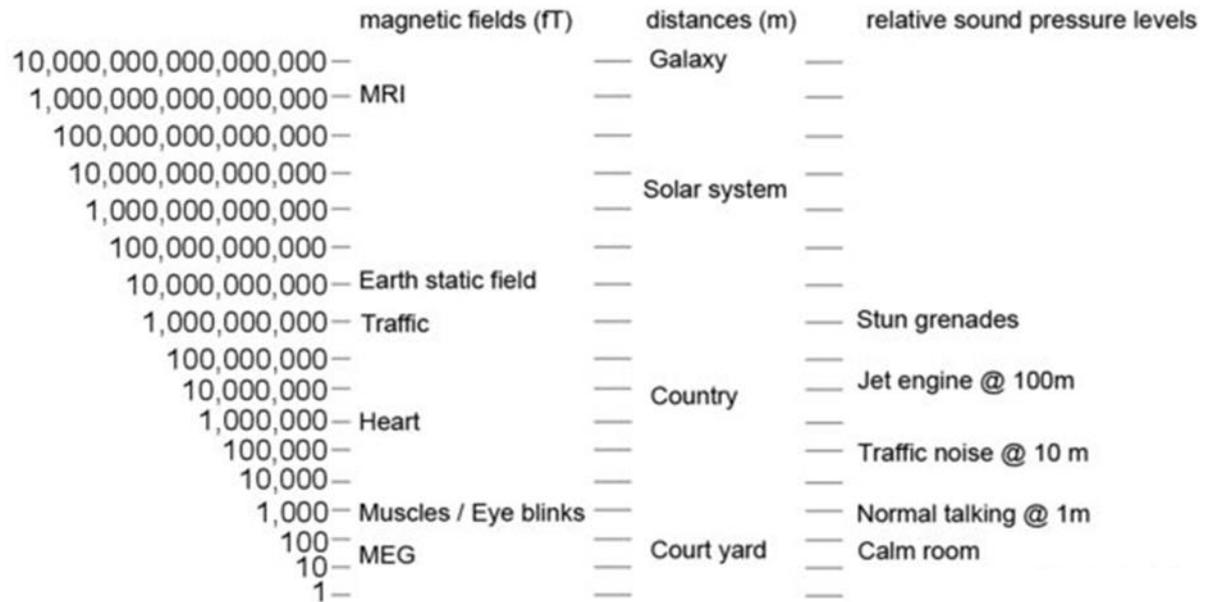


Figure 1.3: Environmental and biomagnetic noise sources.

Scales of magnetic fields relative to those measured with MEG. Numbers reflect femto-Tesla (fT, 1 fT=10⁻¹⁵T). The fields are compared to sound and pressure levels at different distances. Noise reduction is critical because most environmental and physiological noise sources are larger than brain responses; sometimes 10-12 orders of magnitude larger. Note that close proximity to MRI suites introduces immense magnetic fields and the attenuation hereof is therefore critical for MEG recordings in many hospital or neuroimaging laboratory settings. Image adapted from Sylvain Baillet.

Modern noise reduction set-ups consist of four main components:

First, the external magnetic fields are reduced by carrying out recordings inside a magnetically shielded room (MSR). This provides passive shielding against magnetic noise from the environment. High frequencies are attenuated by eddy currents whereas lower frequencies are directed around the shielded room. The external magnetic fields bend around the MSR and thus the noise inside of it is minimized to 10-20 nT. The shielding consists of concentric layers of mu-metal (a nickel-iron alloy) and aluminium.

Second, environmental sources of noise are minimized by placing objects which could interfere with the signal outside of the MSR, and/or only using objects which are guaranteed to not introduce artefacts inside the MSR. For example, the projector used

to present visual stimuli is located outside of the MSR, whereas devices used to record participant's responses, such as button boxes, contain no metallic moving parts.

Third, physiological responses such as those from muscle are reduced as much as possible during recording, and can also, to an extent, be subsequently removed from the data before analysis. These can arise when subjects make saccadic eye movements, blink, swallow, cough, adjust their head position, or when their heart beats. To minimize the saccadic eye movements, subjects are often instructed to look at a fixation cross during a baseline period (although this is far from problem-removing if subjects move their eyes in a way that is stereotypical during the task). To minimize head movement, subjects are instructed to hold still as much as possible, and/or to rest the back of their head on the back part of the inside of the helmet-like structure which contains the sensors (the dewar, **Figure 1.4A**). However, it is still standard to record upwards of 5 mm of head movement, even during short recording sessions (Whalen et al., 2008). This problem is returned to and a possible solution is presented in Chapter 3 where flexible and subject-specific head-casts are described. The cardiac related fields are most often removed from the data offline, through the use of an algorithm that identifies the stereotypical waveform of the heartbeat (the combination of three graphical deflections, the "QRS complex"), and subtracts the wave deflections from the signal.

Finally, reference channels located inside the MEG scanner can be used to measure the ambient magnetic noise. In combination with the signals picked up by the pickup coils and SQUIDs, the information from the reference channels can be synchronously processed such that a third-order gradient of environmental noise sources can be calculated synthetically. This higher-order gradiometer formation is a (CTF-system specific) noise reduction technique carried out in real time. A higher-order gradiometer is created by subtracting a pre-calibrated mixture of reference channels from each gradiometer output. This functions as a form of noise subtraction, as the fields recorded at both the standard and reference channels can be assumed to be environmental, as the fields from the brain will rapidly attenuate with distance from the brain and therefore only be detected by the nearby standard sensors.

Summary

The weakness of the magnetic signals generated by the brain means that reducing externally generated magnetic fields and using highly sensitive devices are

prerequisites for detection. This can be achieved at two levels. First, during acquisition, one can use magnetic shielding, remove ferromagnetic objects from the shielded room, minimize physiological interference from the subject's body (or remove well-characterized events from the data synthetically), and create higher-order gradiometers. In addition, some sensor configurations allow more noise attenuation than others (but this may be a tradeoff with depth sensitivity, see next section).

Second, acquisition protocols and experimental designs can be optimised with respect to minimizing noise (or equivalently maximising SNR) through increasing the number of trials, minimizing blinking, movement, co-registration error, muscle strain etc, and decreasing the brain-sensor distance. To explore these effects further, we now turn to the two types of magnetic sensor that have been employed in the experiments reported in this thesis: superconducting quantum inference devices (or SQUIDs) and optically pumped magnetometers (OPMs).

Superconducting Quantum Interference Devices; SQUIDs

Superconducting Quantum Interference Devices (SQUIDs) rely on the current generated across two Josephson Junctions (JJs) in a superconducting loop. This effect was named the Josephson effect after it was first described by Josephson (Josephson, 1962), giving way for MEG to become a neuroimaging modality ten years later (Cohen, 1972). Modern SQUID sensors use a superconducting loop which is usually made of niobium cooled to superconducting temperatures (**Figure 1.4A,B**). Such temperatures are defined as within 20 degrees of absolute zero, whereas 'cryogenic' temperature refers to temperatures <150 °C. The term cryogenic is therefore used to describe the MEG system's cooling which uses liquid helium to maintain a temperature of ~ -270 °C.

Direct current (dc) SQUIDs output the voltage across the JJs. This voltage can be used to measure the magnetic flux which passes through the loop because they co-vary periodically.

The magnetic flux is measured using a superconducting flux transformer (also called a pickup coil, **Figure 1.4C**). A first-order gradiometer pickup coil consists of two opposite wound wires located some distance apart and both perpendicularly to the surface of the head (**Figure 1.4C**). The advantage of this configuration over simpler ones (e.g. single loops of wire), is that the coil is insensitive to homogenous magnetic

fields such as those created by environmental noise sources. Moreover, the directional sensitivity can be determined by orientating the two loops: *planar* first-order gradiometers are maximally sensitive to the spatial gradient along a particular plane or direction, whereas *axial* gradiometers (**Figure 1.4C**) are maximally sensitive to the magnetic activity which is perpendicular to the surface of the scalp. This thesis uses data recorded and simulated using axial gradiometer dcSQUIDs.

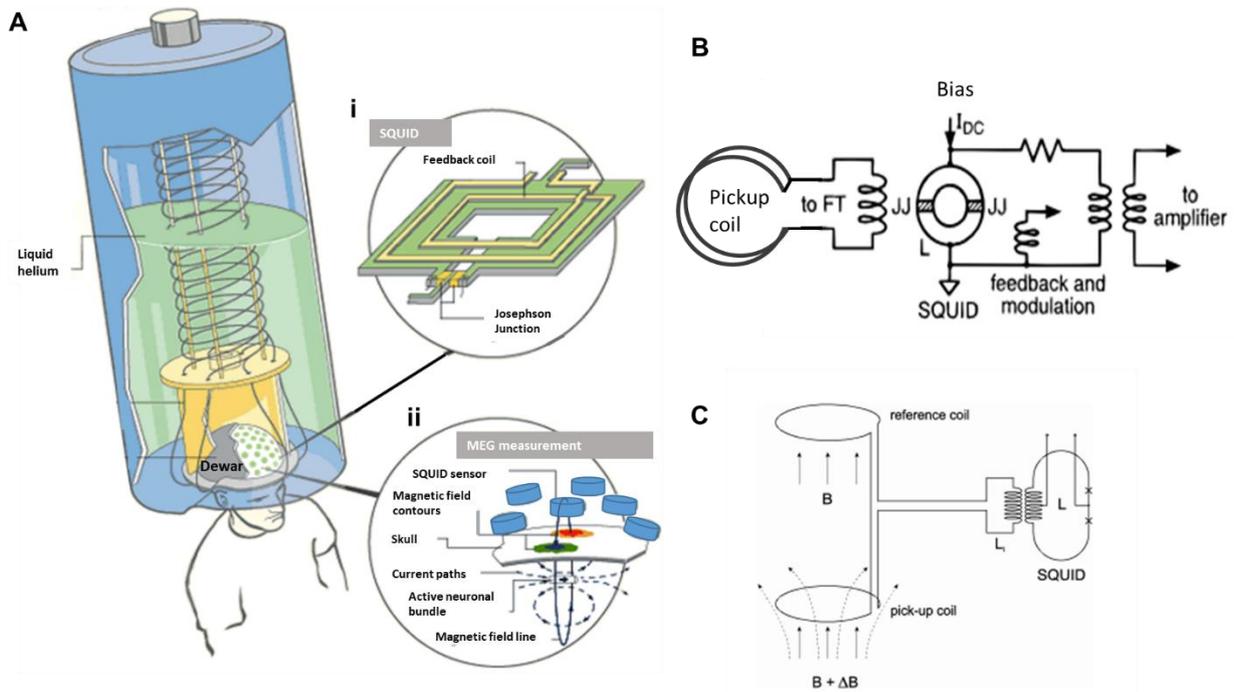


Figure 1.4: MEG system set-up and dcSQUID sensor circuitry.

A) MEG sensor instrumentation. A large tank of liquid helium is used to lower the temperature of the SQUIDs to $-270\text{ }^{\circ}\text{C}$ to enable superconductivity of niobium. i) Diagram of SQUID; Josephson Junction acts as insulator separating two superconductive loops but allowing current to flow between them. ii) Configuration of subset of SQUIDs used to detect magnetic fields produced by neural activity. The magnetic contours represent the in- and out-flow of magnetic fields produced by sources tangential to the scalp. The contours can be detected at a distance and characterized spatially through the use of several SQUIDs. Adapted from (Fishbine, 2003). **B)** Diagram of a dcSQUID. A flux transformer applies magnetic flux which produces oscillations in the SQUID that can be detected (by an external circuit) and amplified. FT = Flux transformer, JJ = Josephson junction, I_{DC} = externally applied direct current, L = inductor of the SQUID. Adapted from (Andra and Nowak, 1998) **C)** Axial gradiometer flux transformer seen from the side. The use of two aligned magnetometers allows substantial noise reduction as distant fields are detected by both and can be subtracted out, while neutrally

generated fields are detected with more strength in the pick-up coil. L_i = inductance of the input coil, L = inductance of the SQUID. From (Andra and Nowak, 1998).

SQUID sensors are most often arranged in a helmet-like structure or a dewar (**Figure 1.4A**), The dewar contains hundreds of sensors distributed across its inner surface as close to the scalp as possible given the requirement for cryogenic cooling. Practically, MEG recording is silent, passive, non-invasive and gives exceptionally high time resolution. The empirical study in this thesis uses a Canadian Thin Films (CTF) system containing 275 sensors, with an axial gradiometer at each location. This system also contains 29 reference channels which can be used to synthesize third order gradiometers, and/or first regress out ambient fields in the case of optically pumped magnetometer recordings.

Until very recently, SQUIDs have remained unchallenged as MEG sensors. A series of combined recent developments in atomic physics and miniaturization however have led to the introduction of a potential replacement technology;

Optically Pumped Magnetometers; OPMs

Optically Pumped Magnetometers (OPMs) are a new generation of MEG sensors which do not require cryogenic cooling. Instead of superconductivity, OPMs rely on the manipulation and interrogation of electron spin in alkali vapour. This section serves as an overview of the physics and mechanics of these new sensors, and compares them to SQUIDs for context.

Similar to SQUIDs, the development of OPMs began close to 50 years ago (Dupont-Roc et al., 1969) but initially had dramatically larger size, large power consumption, as well as poorer sensitivity (note that sensitivity scales with size). Particularly due to their size, these magnetometers could not be used for multi-channel recordings, making them less attractive for MEG experiments (Polyakov, 2003). Over the past decade however, these problems have been solved and OPMs now represent a candidate technology for surpassing and replacing SQUIDs. The primary reason is that OPMs have been miniaturized and operate without cryogenic cooling, meaning that they can be placed directly on the scalp. Critically, they now also provide equal sensitivity to magnetic fields as SQUIDs do (Shah and Wakai, 2013).

Magnetic field sensing by OPMs functions by aligning and then detecting changes in electron-spin precession of vaporous alkali atoms. This is done in a low magnetic field to ensure that ambient external magnetic fields do not influence the spin precession and distort the measurement. First, circularly-polarized light is applied to alkali vapour contained in a glass cell, which causes its single valence electrons to orient their spin along the direction of the polarized light, absorbing the spin of the photons. This process is known as optical pumping. For the spins to orient most efficiently, alkali metals with a single (and thus more easily perturbed) unpaired valence electron such as potassium, caesium, or rubidium are used (**Figure 1.5A**). It is the electron's spin which enables operation of OPMs as it can be used to detect the presence and direction of an external magnetic field. This is because the electron spin precesses around a magnetic field at the precessional or Larmor frequency (**Figure 1.5B**). This frequency refers to the rate of precession of the dipole of the electron around the external magnetic field and therefore reflects the strength of the external magnetic field. If there is no external magnetic field applied, the spins do not change. If on other hand there is, e.g. one generated by the brain, then the spins change, enabling them to be used for detection of the external field. Thus, large polarization can be produced in these electrons through optical pumping with circularly polarized light (**Figure 1.5C**), which can be conceptualised as a baseline condition relative to which the effects of external magnetic field fluctuations can be measured.

“Pumping” refers to the transfer of spins from the light photons in the circularly polarized light, to the valence electrons in the gas. This is a highly effective process, meaning that the spins align to near unity. After spin polarization, the electron spins can be used to detect external magnetic fields by using a probe light (**Figure 1.5D**). The probe light is *linearly* polarised to near resonant for zero applied magnetic field (recall that the nuclear magnetic resonance depends on the magnetic field), and its absorption after passing through the vapour cell can be used to characterize the external field (**Figure 1.5E**). This process is very simple as the changes in light absorption can be quantified based on measurements from a photodiode. The key feature underpinning the relationship between magnetic fields and electron spins is that due to nuclear magnetic resonance (NMR) properties, the resonant frequency of a given substance is directly proportional to the strength of the applied magnetic field. Thus, the resonant frequency can be used to infer the strength of the magnetic field.

The stronger the external magnetic field, the more spin precession and the more absorption and less transmission of probe light. As the spin polarization *difference* in the probe light photons that pass through the cell reflects the presence and magnitude

of an external magnetic field, the probe light passes through the cell uninterrupted and the transmission is therefore maximal when there is no external field (**Figure 1.5E**, black line). The sensitivity of this detection depends on the ambient field being very small or non-existent, as the transmission will be non-informative if it is too far from zero. Moreover, an oscillating magnetic field is applied perpendicularly to the probe light and used to measure the local field's orientation. This applied field is detuned from the gas' resonance (at zero field) so that it does not interfere with the magnitude measurements. The demodulation signal causes the polarization angle to rotate in a known manner and the degree of rotation subsequently measured by the photodiode reflects how well- or mis-aligned the electrons are with respect to the pump beam in a certain direction, which in turn reflects the direction of the external field. The demodulation signal uses a polarimeter, and OPM direction measurements thus work by inferring the orientation of the modulated electron spin precession angle and thereby orientation of the magnetic field from the phase of this signal (**Figure 1.5E**, grey line).

In most OPM devices, rubidium (Rb) vapour is used due to its atomic structure which contains a single valence electron, and its high atomic density and therefore relatively high SNR at low temperatures. For optimal atomic density, the cells are heated to ~ 150 °C. Notably however, the cell vapour-containing cell is very small ($\sim 1\text{cm}^3$) and can therefore be buffered by air such that the surface of the sensor can be room temperature even a few mm away from the sensitive volume. This enables OPMs to be placed directly on the scalp which is the basis of the 5-10 fold SNR improvement (Boto et al., 2016).

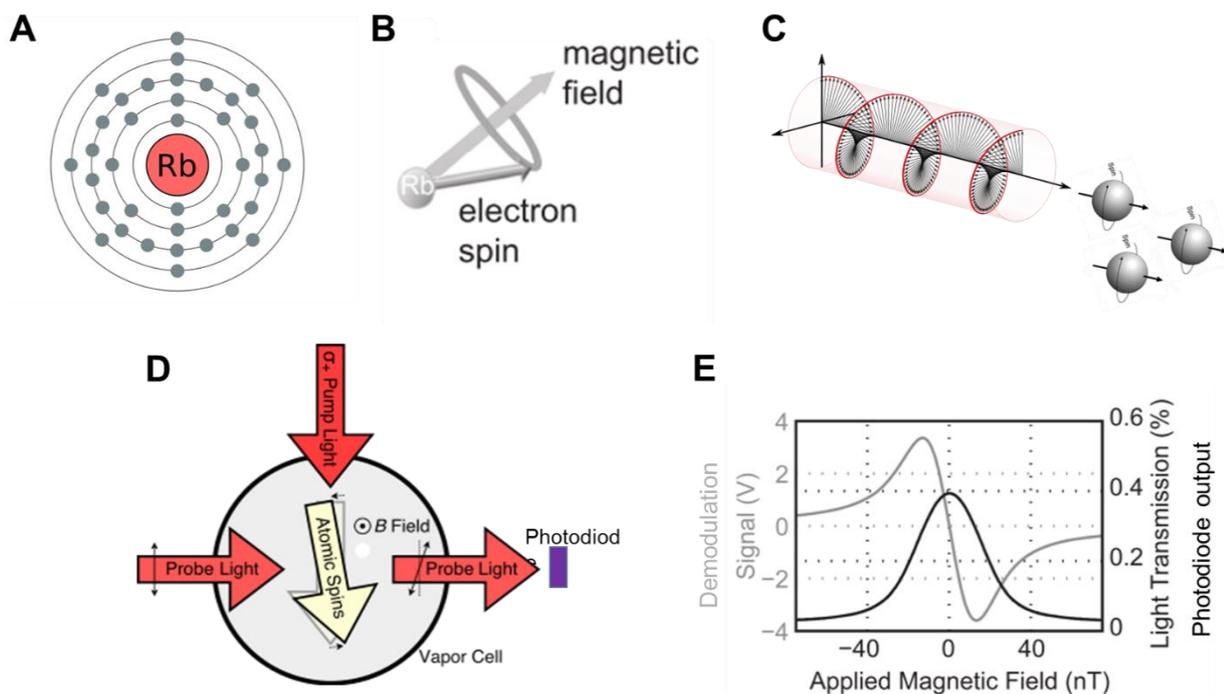


Figure 1.5 OPM principles of operation

A) Electron structure of Rubidium, showing the single valence electron in the 5th and outermost shell. **B)** Relationship between electron spin and magnetic field. The spin of the electron precesses around the magnetic field, be it from circularly polarized light applied by the OPM laser, or neurally generated fields. **C)** Optical pumping. Photons from circularly polarized light transfer their spin to Rb e⁻ which are thus pumped to near-resonance. **D)** Diagram of an OPM. The spin orientations depend on the pump light $\sigma+$ and the magnetic field B . Because the pump light aligns the spins along a known direction, spin divergence from this (near) resonance will only occur if an external magnetic field is present. The probe light is used to detect this. If there is no field, it passes through the cell with maximal intensity and is detected by the photodiode. **E)** Photodiode output and demodulation signal curves. The light transmission, black curve, reveals the presence and strength of an external B field. An oscillating magnetic field is applied perpendicularly to the probe light. This means that the transmitted light is demodulated such that the phase can be used to read out the direction of the field, grey curve.

Until 40 years ago, the largest limiting factor in obtaining higher sensitivity was the simple fact that when atoms collide, they can lose their spin orientation, a process called ‘relaxation’. As collisions cause the electrons to transition into an alternative hyperfine state, it orients and starts to precess in the opposite direction from the rest of the group of atoms. This precession causes decoherence of spins across an ensemble of atoms, which in turn causes the signal to be attenuated. Relaxation can be eliminated by ensuring that the collision frequency is higher than the Larmor frequency of spin precession. This effectively means that the spins do not have time

to precess and decohere between collisions before they collide again and transition back to their original hyperfine state. The high collision frequency is achieved by having at high atom density and high temperature (150-180°C). Thus, the so-called Spin-Exchange Relaxation-Free (SERF) regime has contributed to the improvements in OPM sensitivity by removing the collision-induced sensitivity constraint (Happer and Tang, 1973; Kominis et al., 2003).

Although the introduction of the SERF regime was accompanied by a large increase in sensitive, it is interesting and perhaps counterintuitive to observe that this sensitivity has decreased over time. The explanation lies in the requirements imposed by miniaturization of the sensors. While larger earlier cells were pushed to give extremely high sensitivity, they were also clumsy, impractical, and inherently limited with respect to spatial resolution, even in multi-channel systems. In order for them to be useful for experimental or clinical purposes however, the sensitive volume was reduced and the 'chip-scale' magnetometers were introduced (for example, (Sander et al., 2012)). This reduced volume also reduced the price but limited the sensitivity which nonetheless matches that of SQUIDs. Thus, the SERF regime is central to the recent and dramatic improvements in OPM sensitivity, it can be implemented in small OPMs, but they currently limit the bandwidth to ~100-150 Hz (Sander et al., 2012; Shah and Wakai, 2013). Fortunately however, most brain activity of interest falls within the 0-150 Hz frequency band (Hämäläinen et al., 1993). This means that OPMs are now small and therefore flexible, can be placed close to the scalp, and with a sufficiently large array, can be used to localise and reconstruct neural activity in 3D - a feat which was not possible earlier with larger sensors. Next we discuss and quantify more explicitly how the two types of sensors compare.

Advantages and Disadvantages of OPMs versus SQUIDs

The increased sensitivity of OPMs come at a price. First, the dynamic range of OPMs is limited. The dynamic range is defined as the ratio between the largest and smallest possible measurable values of a changeable quantify, in this case the maximal and minimal magnetic field strength. This is because in very small magnetic fields, the spin precession and continuous pumping interact and static reorientation of the spins occurs (Griffith et al., 2010). Further, OPMs have a lower bandwidth (difference between upper and lower measurable frequency bounds) compared to SQUIDs: ~100-150 Hz for OPMs versus 10,000 Hz for SQUIDs.

Nonetheless, OPMs possess several important advantages over SQUIDs. The most important is that they do not require cryogenic cooling and therefore enable direct and flexible placement of the sensors on the scalp. It follows from the inverse square law, which states that the intensity (here of a magnetic field) is proportional to $1/\text{distance}^2$, that halving the distance between a given source and a sensor equates to quadrupling the signal amplitude. Thus, depending on the distance of the sources from the scalp (larger improvements for superficial sources) and assuming equal noise floors for SQUIDs and OPMs, the source-level SNR improvements will be of a factor around 4 (Boto et al., 2016).

This is promising for the future of MEG research, which may benefit from these SNR improvements in a multitude of ways. OPMs will enable a neuroimaging modality to combine high temporal and spatial resolution. It will also enable flexible sensor configurations such that specific structures can be targeted on a subject-specific basis, be it for basic research to characterize hard-to-access structures such as the cerebellum or brain stem, or for clinical purposes such as localisation of suspected epileptic foci, e.g. in the hippocampus. Moreover, the sensors will be particularly useful for studies involving children that currently can only be scanned with relatively low SNR due to the fixed sensor configuration of SQUID systems.

In particular, the OPMs will potentially be able to compliment and/or replace EEG and intracranial electrode placement prior to epilepsy surgery (Alem et al., 2014), as they can be worn for extended periods of time and thus likely detect infrequent epileptic activity (unlike SQUID-based MEG), while giving superior spatial resolution to EEG and removing the need for dangerous invasive surgery. In addition to these advantages, the acquisition and maintenance prices for these devices are significantly lower than helium-dependent SQUID systems.

As discussed in relation to the SERF regime, another important comparison is with regards to sensitivity. Sensitivity is defined as the ratio of the change in the measurement and the corresponding (true) change in the value of the quantity being measured. In MEG, this is quantified as $\text{fT}/\sqrt{\text{Hz}}$. The sensitivity can also be thought of as the detection noise, i.e. how much uncertainty is included in the measurement. The sensitivity is in part dependent on the size of the sensitive cell and density of the atoms; the more atoms, the higher the sensitivity. However, it is desirable to have small sensors for MEG as this gives way for higher spatial resolution and many-channel measurements. Empirically the sensitivity of OPMs has been measured to be $0.54 \text{ fT}/\sqrt{\text{Hz}}$ with a small (few cm^3) active volume (Kominis et al., 2003). Similarly, a

sensitivity of $0.2 \text{ fT}/\sqrt{\text{Hz}}$ has been reported (Dang et al., 2010) at least over a narrow 1 Hz frequency band. Around the same time, a sensitivity of $5 \text{ fT}/\sqrt{\text{Hz}}$ was reported for an OPM with a sensitive volume of only 1 mm^3 (Griffith et al., 2010). In a shielded environment, the improvement of OPM over SQUID sensitivity has been estimated to be a factor of 10^2 : $\sim 3 \text{ fT}/\sqrt{\text{Hz}}$ for SQUIDs versus the theoretical sensitivity limit of $\sim 0.01 \text{ fT}/\sqrt{\text{Hz}}$ for Potassium (as opposed to Rubidium)-based OPMs. Potassium has the lowest molecule-to-molecule spin relaxation (Allred et al., 2002) and therefore higher sensitivity. However, it requires higher temperatures for optimal atomic density and Rubidium is therefore preferred for MEG applications (this may change if the Potassium cells can be made smaller or more compact however).

Another major difference between SQUIDs and OPMs is the configuration of the sensor. While SQUIDs in the CTF system for example uses axial gradiometers as flux transformers, OPMs are magnetometers which have only a single sensing region and thus no equivalent built-in mechanism for noise reduction (although newer sensors have both field zeroing and modulation coils inside the sensor). Axial gradiometers and magnetometers are sensitive to dipoles in different locations and orientations.

Therefore, while OPMs have a 'zero sensitivity line' directly underneath the magnetometer sensor and thus have maximum sensitivity when sources are positioned off-centre, axial gradiometer SQUIDs are maximally sensitive directly beneath the sensor. Interestingly, the sensitivity changes differ over space between the two configurations as well: the sensitivity decreases more rapidly with distance for axial gradiometers, making them more sensitive to superficial brain sources and less sensitive to deeper sources than magnetometers. This increased sensitivity to deeper sources with magnetometers comes at the cost of needing more accurate models for source reconstruction. This is discussed in more detail in Chapter 5. The main point here is that it is not straightforward to compare the two sensor types directly as they output different aspects of the signal.

Summary

In summary, OPMs work by manipulating and probing well-controlled atomic ensembles inside vapour cells based on the influence of external magnetic fields on electron spin resonance. The most exciting aspects of the sensors are the freedom from cryogenic cooling which enables flexible and direct placement of the sensors on the scalp, improving the SNR at least 5-fold. More recently, the miniaturization of these

sensors has enabled them to be configured into multiple channel set-ups, which enable large improvements in spatial resolution due to the increased sensitivity compared to SQUIDs. This development has created a final push towards potentially adopting OPMs as a new technology for MEG. However, OPMs have bandwidth compared to SQUIDs, but these are still sufficient for recording brain activity. The dynamic range is smaller but can be extended by using local feedback coils to cancel out external fields.

In relation to the main aims of this thesis, the current magnetometer set-up of OPMs, although likely to soon develop to gradiometers for noise reduction purposes, will most likely be more optimal than SQUIDs for detect hippocampal sources due to their increased SNR from being closer to the source. This notion and other anticipated future developments will be revisited in the discussion.

MEG analysis techniques: theory and methods

An externally measured electromagnetic field can be used to estimate neural activity in terms of its three-dimensional source configuration and time course. This requires two steps: construction of a forward model which describes the predicted scalp distribution for a given source with a specific orientation, magnitude and location, and subsequent inversion of this forward model to estimate the most likely spatial configuration of sources giving rise to the measured signal. All the inversion methods used here are parametric empirical Bayes linear inverse methods - the definitions and implications thereof will be described in this section. All source reconstruction methods rely on carrying out the following steps: preprocessing including filtering and removal of possible artefacts present in the data, source space modelling, data co-registration, forward computation, before finally carrying out the inverse reconstruction. Here we focus on the latter four. The methods for inverse reconstruction vary with respect to assumptions about covariance among sources. All simulations and analyses were carried out using the Statistical Parametric Mapping (SPM) software version 12 <http://www.fil.ion.ucl.ac.uk/spm/>.

This section serves to introduce in more detail the analysis methods and theoretical considerations underpinning source reconstruction methods used in this thesis. First, the forward and inverse modelling will be discussed, data co-registration will be introduced, and the general linear model will be described. This model proposes a solution to the inverse problem and as such comprises the linking principle between observed data and source estimates. Next, the Bayesian implementation of the inverse methods applied in this thesis will be discussed. Following this, a brief overview of the commonalities and mathematical terms of the different methods will be given (see (Belardinelli et al., 2012; López et al., 2014; Wipf and Nagarajan, 2009) for detailed descriptions). Finally, the Free energy principle and Bayesian model comparison will be considered in the context of source reconstruction. The objective in later chapters is to set up a framework for direct comparison of competing generative (forward) models, and examine this comparison across a range of reconstruction methods, empirical factors, and subjects.

Forward Modelling and Data Co-registration

The 'forward' model is a description of the primary current distribution and its propagation through the head before reaching the sensors (Equations 1.1-1.3). It relates the measured sensor distribution with dipoles with known location, orientation and magnitude.

Importantly therefore, the forward problem that the forward model solves is well-posed. This is because there is only one solution for each current dipole, as the laws of electromagnetism are linear. It follows that if the primary current density consists of N terms, then each measurement can be decomposed into the sum of N terms where each represents the part of the measurement that is generated by a single potential source. In order to compute this forward model, one must first specify a generative model of the brain, scalp and skull (where the first is the most important in MEG). This generative model is informed by and based on neurophysiology, and it is possible to incorporate unique anatomical characteristics such as shapes of certain structures. The latter is used in this thesis for incorporating the surface of the hippocampus into the generative model, and evaluating whether this improves the ability of the generative model to explain the data parsimoniously. More commonly and generally, the forward model is specified by a manifold of the cortical surface based on an anatomical MRI image. Specifically, the cortex, skull and scalp surfaces can be extracted from T1-weighted anatomical scans using automated segmentation tools such as Freesurfer (Dale and Sereno, 1993; Fischl, 2012). In turn, this gives way for explicit modelling of the dipole orientation, which is usually specified to be orthogonal to the cortical surface mesh where each vertex constitutes a putative source. The main advantage of this approach is that given fixed dipoles, the electromagnetic forward model can linearly map each source onto each sensor based on approximations to Maxwell's equations. Accurate solutions describing the signal can be obtained using methods such as the boundary element method (Brebbia and Dominguez, 1989) which incorporates geometric representations of each surface.

Accurate forward modelling requires accurate estimation of the location of the brain in order to be useful. Co-registration is the process by which the functional MEG data are aligned or co-registered with the structural generative model (based on MRI data). However, if the co-registration is inaccurate and contains 5 or more millimetres of error, or there is an equivalent amount of (unmodelled) head movement during scanning, then it is reasonable to instead use a non-linear registration of the subject's MRI scan to a canonical template brain (Henson et al., 2009; Mattout et al., 2007; Troebinger et al., 2014b).

The Inverse Problem

Once a forward model has been constructed, it is inverted such that the underlying neural activity generated by the observed data can be estimated. Unlike the forward problem, the inverse problem is mathematically ill-posed, as there are effectively an infinite number of possible inverse solutions to a given forward model. In other words, even if the experimenter knew the exact magnetic field at all points outside the head, s/he would still be unable to determine with certainty the configuration of sources inside the head. The problem can be compared to estimating the configuration of hands and fingers based on a shadow (**Figure 1.5a**). Practically however, this limitation can be overcome if one is willing to make some simplifying assumptions to constrain the solution space. These are often already present and well-known, e.g. approximate relative sizes of objects in relation to distance (**Figure 1.5b**) or, in MEG, brain structure and function. In order to find a unique solution, one must specify a set of prior assumptions to constrain the inverse solution (Baillet et al., 2001). For MEG inverse solutions, these constraints can be anatomical, functional, and/or mathematical. In essence, the constraints or ‘priors’ act to define the solution space such that the most probable solution (given the priors) can be identified. The most likely priors are therefore those that maximise the model evidence for a given MEG dataset (Friston et al., 2008a; López et al., 2014; Wipf and Nagarajan, 2009). Hence, the Bayesian framework allows one to ask which model or set of priors is the most likely, given the data at hand. The approximation of the model evidence will be discussed in the *Free Energy* section below.

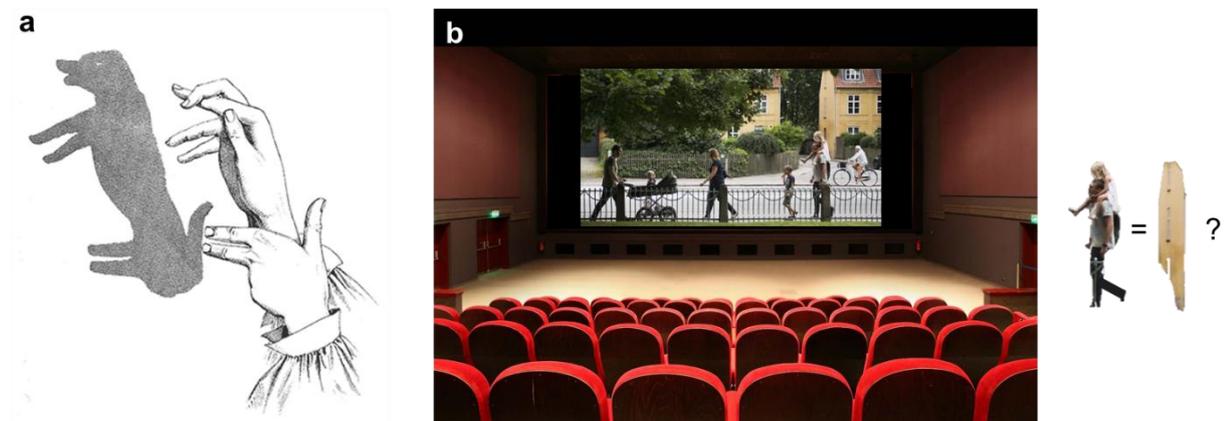


Figure 1.5: Illustration of ill-posed nature of the inverse problem.

a) Non-uniqueness of inverse problems: a 2D shadow can easily be computed given a source of light and a surface (akin to solving the forward problem). Conversely, describing a 3D configuration of hands and fingers based on this shadow is under-determined (akin to the inverse problem); many configurations could give rise to the same shadow figure. The challenge for the ill-posed inverse problem in MEG is reconstructing 3D neural source configurations from a lower-dimensional projection. Image adapted from *Hand Shadows to be Thrown Upon the Wall* (Henry Bursill, 1895, <http://www.gutenberg.net>). **b)** More realistic version of how the inverse problem is solved in an everyday context: by looking at the image on the screen, it appears that the man carrying the girl on his shoulders are the same height as the yellow house. Given a priori knowledge about the relative sizes of humans and houses however, it is easy to explain away this apparent equality with distance between the two.

A well-validated strategy for solving the inverse problem is the distributed approach, where a large number of fixed dipoles fill the solution space and only their amplitudes (and not orientations or locations) are estimated based on the data. This allows an algebraic solution to the inverse problem (Grech et al., 2008).

The Generalized Linear Model

A distributed set of neural sources that linearly map onto sensors placed outside the head are employed by all the reconstruction algorithms used here (Hämäläinen et al., 1993):

$$Y = LJ + \varepsilon \quad (1.4)$$

Where the dataset $Y \in \mathbb{R}^{N_c \times N_t}$ contains information of N_c sensors at N_t time points, the lead field $L \in \mathbb{R}^{N_c \times N_d}$ incorporates our assumptions about cortical folding and head location between N_c sensors and N_d sources, and $J \in \mathbb{R}^{N_d \times N_t}$ is a matrix containing N_d amplitudes by N_t time points of unknown primary current density parameters (i.e. neural sources). ε is a zero-mean Gaussian noise distribution which incorporates both sensor noise and uncertainty about the propagation through L . Note that L is defined as the propagation model of an MEG signal that is produced by a source of unitary strength, and is completely determined by the sensor configuration and volume conductor physics, as described in terms of the forward model. This function also contains all known details about the measurement set-up and physical properties of the brain. Since these are unlikely to change over the course of the experiment, the lead fields is only computed once per dataset. This is valid provided that the location

of the head relative to the sensors is unchanged – whether or not this assumption is valid will be discussed later.

The generalized linear inverse solution describes the neuronal activity, $J(t)$ at any latency or time point t . Because the number of potential sources greatly exceeds the number of sensors ($N_d \gg N_c$), L is ill-posed and therefore cannot be inverted directly. This problem is simplified by assuming that the source amplitudes J are a zero-mean Gaussian distribution with covariance $\text{cov}(J) = Q$, which gives the generalized linear inverse solution:

$$\hat{J} = QL^T(Q_\varepsilon + LQL^T)^{-1}Y \quad (1.5)$$

Where Q_ε describes the sensor-level covariance $\text{cov}(Y)$ and $\hat{J} \in \mathbb{R}^{N_d \times N_t}$ contains the estimated N_d source amplitudes by N_t time points. This general expression is used across most distributed source reconstruction algorithms which employ Gaussian assumptions. Given that Y is known and L can be computed from the head model (and sensor configuration and volume conduction principles), the only parameters needed to compute the source current estimates in \hat{J} are estimates of the sensor and source level covariance matrixes, Q_ε and Q . The differences between inverse schemes arise from how Q – the source level covariance - is defined, and this will be described in the section *Covariance Matrix Specification Using Different Functional Priors*.

Thus, having an algebraic solution to the inverse problem implies that the inverse problem can be formulated as a probabilistic generative model of the data. The term probabilistic refers to the joint probability of all variables in the model while the term generative implies that the model describes how the data were generated. This is done by including all possible (and not mutually exclusive) variables and their estimated probabilities in the joint probability, i.e. calculating the combination of factors which could have given rise to the data. These factors include locations and orientations of possible sources and are described in greater detail below. Moreover, the generative model is hierarchical with two parameter levels. Each level has its own prior precision which determines the importance of the prior relative to the data (in a Bayesian context relative to the likelihood). Optimising these precision values as free parameters is therefore equivalent to optimising the balance between the data and the priors at hand (Mattout et al., 2006; Phillips et al., 2002). This is a critical feature because it is effectively empirical model optimisation since the parameters can be estimated from the data.

The two levels are the sensor level which is directly measured, and the source level which is inferred. The source level is higher in the hierarchy and thus its ‘hyperparameters’ represent priors on the lower level sensor parameters. This relationship means that the source level hyperparameters scale the sensor level prior distributions such that the likelihood is maximal given the data. Notably, this method assumes Gaussian distribution of the priors (Belardinelli et al., 2013; Friston et al., 2008b; Henson et al., 2011; Phillips et al., 2005).

Bayesian Implementation

Applying Bayesian methods to the inverse problem is helpful because they enable the solutions (i.e. estimated locations or time courses) to take the form of posterior distributions as opposed to point estimates. A posterior distribution is the distribution of predictions (or unobserved observations) conditional on the observed data. In the context of source reconstruction, the posterior distribution describes the subset of possible sources which are most likely to have given rise to the observed data. Using a Bayesian framework also makes it possible to marginalise out irrelevant variables through integration.

The basis of the Bayesian implementation is that the recorded activity over sensors, Y , is used to estimate the distribution of putative sources in the brain. The inverse solution constraints take the form of prior probabilities of the source activity $p(J)$ and that these priors are informed by anatomy and neurophysiology (and combined with the physical properties of volume conduction). The priors are then used to estimate the posterior probability of the source activity given the data $p(J|Y)$ through combination with, or weighting by, the likelihood of the data $p(Y|J)$. As such, Bayes theorem takes the form (Grech et al., 2008):

$$p(J|Y) = \frac{p(Y|J)p(J)}{p(Y)} \quad (1.6)$$

Where evidence for the recorded data $p(Y)$ is considered to be constant, given a constant dataset. This term also acts as a normalisation term because it is the denominator of the equation. Gaussian refers to the assumption of normally distributed and mean-zeroed data in the time domain. The estimated magnitude of J , \hat{J} can be found by taking the expectation of the posterior:

$$\hat{J} = E[p(J|Y)] \quad (1.7)$$

The covariance of $p(J|Y)$ is equivalent to $\Sigma_J = (L^T Q_\varepsilon^{-1} L + Q^{-1})^{-1}$. As this ($N_d \times N_d$) source covariance is computationally expensive, it is commonly replaced by the modelled sensor covariance (Friston et al., 2008b) which is given by:

$$\Sigma = Q_\varepsilon + LQL^T \quad (1.8)$$

Where Q_ε is the sensor noise covariance, L is the lead field, Q is the prior source covariance matrix and $(.)^T$ denotes the transpose operator. This in turn enables projection of the source space covariance components into the (typically smaller) sensor space. Given the algorithm-specific source covariance component(s) C_i and the covariance estimate of the sensor noise Q_ε , the sensor covariance can be modelled as Σ for optimising the hyperparameters:

$$\Sigma = e^{\lambda_0} Q_\varepsilon + \sum_{i=1}^{N_q} e^{\lambda_i} L C_i L^T \quad (1.9)$$

Where h_i is now expressed as e^{λ_i} which constrains its value to be positive, ensures that the optimisation is convex and that the prior on the hyperparameters follows a Gaussian (this equates to assuming a log-normal distribution on the scale parameters e^{λ}) (Friston et al., 2008a; Wipf and Nagarajan, 2009).

Principles of Inverse Reconstruction

The three different inversion schemes used in this thesis are briefly described here with respect to their commonalities and differences. What differs between them is the choice of putative sources (the prior set) and the accompanying assumptions about source covariance Q . The rationale for each is the same, however, and all three algorithms are PEB algorithms. As for these three schemes, most popular inversion schemes differ only in the choice of the form that the source covariance Q takes (Equation 1.5) (Friston et al., 2008a; Mosher et al., 2003; Wipf and Nagarajan, 2009) - that is, these three inversion schemes differ in their choices of source covariance Q .

The sensor noise is assumed to be identically independently distributed (IID), i.e., an identity matrix scaled by the so called regularization parameter h_0 . This means that

the same noise level is assumed on all channels such that $\text{cov}(\varepsilon) = Q_\varepsilon$. It can be written as:

$$Q_\varepsilon = h_0 I_{N_c} \quad (1.10)$$

Where $I_{N_c} \in \mathbb{R}^{N_c \times N_c}$ is an identity matrix, and h_0 is the sensor noise variance. Implicit in this formulation is the assumption of uniformity (that the noise variance is the same across all sensors). h_0 can also be informed by empirical recordings of an empty MSR to approximate the true sensor-level covariance (Henson et al., 2011).

Another important parameter common across the inversion schemes employed here is smoothness, the spatial extent of each source prior. We include this because it is known a priori that neuronal currents display local coherence. The smoothing function used to determine the modelled smoothness is computed locally on the vertices of the anatomical subject-specific MRI-derived surface mesh. We use a Green's function based on a graph Laplacian proposed by (Harrison et al., 2007) which can be described as:

$$G(s) = \sum_{i=0}^8 \frac{s^i}{i!} A^i \quad (1.11)$$

Where the adjacency matrix A denotes the neighbourhood properties of the vertices where $A = 1$ if there is face connectivity and $A = 0$ otherwise. The smoothness parameter s determines how far from the central vertex the G function connects the patch points. In SPM, the default value, and the value we use here is $s = 0.6$. This provides a trade-off between spatial accuracy and local coherence, it corresponds to assuming an effective local coherence or patch diameter of approximately 10 mm if the mesh density is approximately the same as a "normal" mesh with 8196 vertices.

Covariance Matrix Specification Using Different Functional Priors

We now turn to the differences between the inversion schemes utilised here: Minimum Norm Estimation (MNE), Empirical Bayes Beamformer (EBB) and Multiple Sparse Priors (MSP).

Bayesian MEG inversion algorithms differ only with respect to definition of the prior source covariance matrix Q (Mosher et al., 2003; Wipf and Nagarajan, 2009). Here we

briefly describe the differences between Q s specified using Minimum Norm Estimation (MNE), Empirical Bayes Beamformer (EBB) and Multiple Sparse Priors (MSP).

The standard minimum norm estimation (MNE) represents the simplest assumption about prior source covariance Q , namely that the sources are independent and identically distributed. This means that they have the same variance and no covariance and can therefore be described by the covariance matrix:

$$Q_{MNE} = I \quad (1.12)$$

Where $I \in \mathbb{R}^{N_c \times N_c}$ is an N_c sensors by N_c sensors identity matrix.

The Empirical Bayes Beamformer (EBB) inversion scheme is similar to MNE as there is a single diagonal source prior covariance matrix which can be written as:

$$Q_{EBB} = \text{diag}(\sigma^2) \quad (1.13)$$

Where σ^2 is the source variance. Unlike MNE, this variance value is estimated directly from the data under the standard beamforming assumption that there are no spatially separated but temporally correlated sources (Belardinelli et al., 2012; Van Veen et al., 1997). Thus, while MNE assumes that all sources are potentially active with equal probability, EBB selects and weights priors based on their putative contribution to the measured signal. Specifically, for every dipolar location θ , the source variance σ^2 is calculated as follows (Hillebrand and Barnes, 2005; Mosher et al., 2003):

$$\sigma_\theta^2 = (L_\theta^T C_b^{-1} L_\theta)^{-1} \quad (1.14)$$

Where $C_b = YY^T$ is the sensor-level covariance matrix and L_θ represents the effective, smoothed lead field for a patch centred at dipole location θ . $(\cdot)^T$ denotes a transpose operator.

The Multiple Sparse Priors (MSP) inversion (Friston et al., 2008a) models a set of sparse local spatial patches (as opposed to a single cortex-wide pattern) and prunes away those patches which do not explain variance. Importantly, MSP is a more general form of the approaches described above because the structure of the prior components in Q can take any other form (if it is more appropriate), including those of EBB and MNE. This is because the prior source covariance is a weighted sum of a set of (multiple sparse) prior components, one per spatial prior: $Q = \{C_1, \dots, C_{N_q}\}$ where N_q denotes the number of priors covering the mesh. The default $N_q = 512$ in SPM. The priors constrain the source space such that the algorithm is forced to explain the data

using the priors given such that a poor model (i.e. a set of incorrect priors) will have a low model evidence. Both generative models used with MSP have 90 identical randomly distributed cortical priors but differ on the inclusion of hippocampal priors (the remaining 10 cortical priors in the cortical model are also randomly distributed across the cortex). The (global) prior source covariance matrix can be expressed as:

$$Q_{MSP} = \sum_{i=1}^{N_q} h_i C_i \quad (1.15)$$

Where each $C_i \in \mathbb{R}^{N_d \times N_d}$ is itself a prior source covariance matrix with each prior component corresponding to a smooth surface patch. These covariance components are individually weighted by hyperparameters $h = \{h_1, \dots, h_{N_q}\}$. The larger a given hyperparameter, the larger the prior variance of the patch. MSP can thereby optimize the hyperparameters so as to best fit the modelled covariance to the data covariance (at sensor-level) by mixing and pruning (hyperparameter down-weighting) the priors such that the model evidence is maximized. Whereas MNE and EBB use only a single hyperparameter to fit the data covariance, MSP uses one per spatial prior. MSP is a generalisation of Bayesian inversion algorithms as it can take any other form if this is optimal for explaining the data at hand parsimoniously.

After implementing the functional prior assumptions as described above, the algorithm-specific Q is empirically optimised whereby that the current density can be estimated and most likely source distribution inferred. The optimisation is based on an approximate Bayesian inversion scheme, Variational Laplace (Friston et al., 2008b), which assumes that the posterior distribution of J ($J \in \mathbb{R}^{N_d \times N_t}$ which describes the amplitude of N_d current dipoles over N_t time samples) is Gaussian. The result is a set of hyperparameters that maximize the model evidence for the given data, and which are used to specify Q in the subsequent data inversion step (see for example Grech et al, 2008 for details).

Free Energy

Each inversion returns a negative variational Free energy value (F) which approximates the model evidence $p(Y|m)$ where Y is the data and m is the model (Friston et al., 2007; Wipf and Nagarajan, 2009). Because F represents a trade-off between complexity and accuracy (Penny, 2012), it is used as the cost function to find

the best hyperparameters when the models underlying the source reconstruction are linear and Gaussian.

The Free energy is used as the objective function to fit the modelled covariance Σ (Equation 1.16) to the actual data covariance Σ_γ . The Free energy, F , is expressed formally (Friston et al., 2007) as

$$F = \frac{N_t}{2} \text{tr}(\Sigma_\gamma \Sigma^{-1}) - \frac{N_t}{2} \log|\Sigma| - \frac{N_t N_c}{2} \log 2\pi - \frac{1}{2} (\hat{\lambda} - v)^T \Pi (\hat{\lambda} - v) \quad (1.16)$$

$$+ \frac{1}{2} \log|\Sigma_\lambda \Pi|$$

Or equivalently, $F = -[\text{model error}] - [\text{size of model covariance}] - [n \text{ Samples}] - [\text{error in hyperparameters}] + [\text{error in covariance of hyperparameters}]$

Where $\text{tr}(\cdot)$ is the trace operator, N_t is the number of time samples or temporal modes (here $N_t = 1$), Σ_γ is the measured data covariance and Σ is the modelled data covariance. There are N_c sensors or spatial modes (here $N_c = 274$). Where λ and v are the prior and posterior means, and Π and Σ_λ the prior and posterior precisions of the hyperparameters, respectively. We use the SPM default values $\lambda = -32$ and $\Pi = 1/256$ which makes the hyperparameters weakly informative.

This optimization can be thought of as a process to minimize the number of source patches but still explain the maximum amount of data. The mixing and pruning of priors means that for large numbers of priors, the optimization can get trapped in local extrema. One practical solution to this is to run the same algorithm many times with different sets of priors (spatial patches) (Troebinger et al., 2014a).

Perhaps the most important principle of the Free energy formulation for the purposes of this thesis is that it can be divided into two constituent components: accuracy and complexity. The accuracy is given by the first three terms in the equation while the complexity (which is what differs from other Bayesian approaches (Wipf and Nagarajan, 2009)) is given by the latter two terms. The accuracy reflects how well the model explains variance in the data while the complexity reflects the error in the approximations of the hyperparameters.

As F approximates the model evidence for a generative model used to invert a set of data, Free Energy values obtained with different models can be used for comparison of the likelihood of these models. In this model comparison framework, properties of the generative model are changed, and the F values obtained for the different parameter values can be compared with respect to how well they describe the data parsimoniously simply by subtracting one F value from the other. It is thereby possible to quantify the difference in (approximated) model evidence which enables hypothesis testing through variations in the generative model. For example, hypotheses could relate to anatomical structures involved in generating the signal. The application of this model comparison approach to MEG data analysis has successfully been demonstrated elsewhere (Henson et al., 2011, 2009; Lopez et al., 2013; López et al., 2014; Penny, 2012; Stevenson et al., 2014; Troebinger et al., 2014a). Because of the Bayesian context, the F value difference is equivalent to calculating a Bayes factor. In line with Bayesian convention, a positive difference means that the first model in the equation is $\frac{1}{1+e^{\Delta F}}$ more likely than the second. A significance threshold is defined at 3 where, because of the log distribution of the Bayes factor of F difference, one model is ~20 times more likely than the other. Critically, model comparison is only valid however when the data is the same and can only be used to infer the relative fitness of two models – not whether, or what form a potentially better one might take.

Summary

The general linear inverse expression describes how sensor-level data modelled as a distribution of primary currents can be inverted such that the locations and time courses of these currents or neural sources can be reconstructed. This is done by constructing a forward model which simulates the field distribution for a current dipole in a given orientation, location, and with a given source strength inside a volume conduction model of the brain.

Because the inverse problem is ill-posed, it is necessary to specify prior constraints. This can be done through prior distributions describing putative sources in a Bayesian context. This gives a posterior distribution over the potential sources by incorporating information about covariance present in the data into the solution, in a two-layer hierarchical model of how the data were generated.

Assumptions about source level covariance are expressed through different inverse reconstruction algorithms. These can also be thought of as functional priors (as

opposed to spatial priors). Within the Parametric Empirical Bayesian framework in SPM, any inverse solution also returns an approximation of the model evidence (a Free energy value) which gives a trade-off between the accuracy and complexity of the solution, and can be used to compare models (different functional and/or anatomical priors) based on the same data.

Hippocampus and hippocampal oscillations

The hippocampus is a small bilateral curved structure constituting part of the archicortex in the medial temporal lobe (**Figure 1.7B**). Its shape is similar to a sea horse from which it also takes its name in Greek. It is ~5 cm long in adult humans (Schultz and Engelhardt, 2014). It is a uniquely important brain structure in humans due to its role in episodic memory. This function is thought to have co-evolved alongside navigation and spatial memory functions such that memory/learning, spatial navigation and planning could be supported by the same neural circuits (Amaral and Witter, 1989). Underpinning hippocampal mechanisms which support episodic memory and spatial navigation are, most prominently studied and thus most well-understood, 4-8 Hz theta oscillations (Vanderwolf, 1969; O'Keefe and Nadel, 1978; Buzsáki, 2005). Hippocampal theta is very well studied in rodents, and possesses a wide range of directly quantifiable relationships between behaviour and features of the oscillatory activity (such as frequency, phase and power). For example, theta frequency and amplitude increases with running speed (McFarland et al., 1975) but the frequency decreases with environmental novelty (Jeewajee et al., 2008).

Overall, we are interested in knowing whether (and how) we can translate the rodent findings to humans and potentially further nuance them in this context. At present, efforts to do so is occasionally possible in epileptic patients but for generalizability, statistical power, and experimental freedom we must make these recordings non-invasive.

Other human brain rhythms are predominantly generated by superficial sources and are therefore relatively easy to measure using MEG. Hippocampal theta on the other hand, is not. Thus, hippocampal theta is less well-studied in humans due to methodological difficulties (which are addressed by this thesis), but iEEG recordings from epileptic patients suggest that there are memory correlates (such as the subsequent memory effect, (Sederberg et al., 2003)) as well as bouts of theta oscillations during movement in virtual reality (VR) environments.

Relationships between Hippocampal Theta Oscillations and Behaviour

In rodents, a large set of complex and interesting relationships between theta and behaviour have been demonstrated. For example, it has been shown that there is a

relationship between theta and movement/arousal (Green and Arduini, 1954), that theta frequency and power correlate with running speed (McFarland et al., 1975), that theta frequency is reduced by anxiolytic drugs (John et al., 2014), that eliminating theta impairs spatial memory function (and gets rid of grid cell firing patterns) (Brandon et al., 2011; Givens, 1995). Two main domains have developed as a result of these studies: one investigates the role of theta in physiological terms, linking synaptic changes to behavioural changes in relation to theta. The other has emphasized the role of theta in computing the location of an animal during spatial navigation.

The first domain which places the main focus on mechanisms related to memory and learning is based on a set of documented links between synaptic potentiation, and theta. For example, it has been shown that there is increased efficiency of memory encoding during periods where theta amplitude is high (Seager et al., 2002), and it has since been suggested that long-term plasticity may be induced by theta. Specifically, it has been proposed that theta supports memory by providing a (timing) signal which causes a population of simultaneously active cells to spike within a short temporal window, in turn causing Hebbian plasticity and thus, long term changes in synaptic connections (Buzsáki, 2005; O'Keefe and Recce, 1993).

The second domain on the other hand has focused on the notion of a cognitive theta map (O'Keefe and Nadel, 1978; Tolman, 1948). The basis of this theory comes from a study which showed that rodents build an internal map of the environment which enables them to find a goal location regardless of starting point – and to take shortcuts, as opposed to simply link actions to stimuli and follow previously executed routes (Tolman, 1948). Importantly, these processes have since been shown to be dependent on the integrity of the hippocampus (Morris et al., 1982). Currently, this field concerns itself with the interactions and computations of the different cell types and functions found in and around the hippocampus. These include (but are far from limited to) hippocampal place cells (O'Keefe, 1976), entorhinal grid cells (Hafting et al., 2005), and accounts of how theta mediates informative combinations of the neural dynamics hereof, e.g. the role of phase in determining where within a place field an animal is currently, or soon to be, located (Burgess and O'Keefe, 2011; Buzsáki and Moser, 2013; Moser et al., 2008). Thus, theta oscillations modulate the activity of hippocampal place cells (O'Keefe and Recce, 1993; Huxter et al., 2003). This branch of research also concerns itself with the different cognitive processes which constitute navigation. These processes include path integration whereby the internal self-motion cues are integrated without reference to external cues, in order to encode the relative

spatial location. One interesting example of the link between path integration of theta showed that theta power is dependent on movement velocity, but that the gain of this function is higher when the animal moves on its own, compared to being moved, or having the external cues move (Terrazas et al., 2005).

In humans, the body of literature is far smaller but supports both theories of hippocampal theta functions: there is evidence that theta power correlates with memory performance (Ekstrom et al., 2005; Lega et al., 2012), and similarly, some evidence that theta is associated with movement (especially movement onset) in virtual reality (VR) environments (Kaplan et al., 2012a). Moreover, theta power (synchrony) has been shown to selectively correlate with retrieval of spatial information that is relevant for navigation (de Araújo et al., 2002). Generally however, in order to begin to address the outstanding questions, as well as test new predictions, a reliable and non-invasive method for recording hippocampal theta is needed (Dalal et al., 2013a; Riggs et al., 2009).

Preoperative epilepsy patients with recording grids in the hippocampus have provided valuable but rare insights into theta dynamics in humans (Jacobs et al., 2013; Lega et al., 2012; Tesche and Karhu, 2000). However, such recordings are suboptimal for ethical and practical reasons, as well as being infrequent and potentially non-representative of healthy hippocampal processing. Thus, a non-invasive and reliable alternative could potentially replace these invasive pre-operative procedures. This could offer a much safer, more comfortable, cheaper and faster method of estimating the epileptic foci, if effective and spatially precise. Further, it could enable a new range of cognitive experiments which could address the role(s) of hippocampal oscillations in the human brain. In addition, it is likely that the paradigms could be more complex and thus representative of real-life navigation than is currently the case with patients who are often drowsy, elderly, unwell, and/or off medication.

In this thesis, the focus is on using and developing new MEG methods for detecting hippocampal activity. Since a large and growing body of both rodent electrophysiology and human neuroimaging work has focused on the ability of the hippocampus to represent and process spatial information, we use a well-validated spatial cognition task to activate hippocampus in the empirical validation of the methods. This section serves as a brief introduction to the relevant spatial cognition literature, overview of anatomical and electrophysiological features incorporated into the MEG forward model, and a discussion of previous MEG experiments showing evidence of hippocampal activity.

Viewpoint Independence

A cognitive map can be defined with respect to the behaviours it allows and a central feature of the hippocampal cognitive map is that it is viewpoint-independent. This is an ideal format of spatial knowledge since it can enable behaviours such as taking a novel short cut, flexibly planning and imagining routes, finding a desired goal location from any starting position, etc.

This form of perspective is called allocentric and can be contrasted with an egocentric or first-person viewpoint (**Figure 1.7A**). Egocentric representations of space have been shown to localise to the parietal lobe, where lesions lead to egocentric spatial processing deficits (Bisiach and Luzzatti, 1978), replicated in fMRI where the posterior parietal cortex for example tracks egocentric direction of goal locations (Spiers and Maguire, 2007). Allocentric spatial memory representations however have been shown to be hippocampal: patients with bilateral hippocampal lesions are strongly impaired on recognition of object arrays only when the viewpoint has been shifted (King et al., 2002), suggesting an impairment specifically in allocentric representation (be it encoding or retrieval or both). Further, the hippocampus has been shown in fMRI to be more active during novel wayfinding than stimulus-based route following (where an allocentric mental representation is required for the former but not latter) (Hartley et al., 2003), and when planning routes during real-time navigation (Spiers and Maguire, 2006). Similarly, the amount of allocentric knowledge acquired in single trials (measured behaviourally through trial-by-trial improvements) correlates with hippocampal activation (Doeller et al., 2008). A classical rodent example of this anatomical dichotomy between allocentric and egocentric processing showed that inactivating hippocampus and striatum makes rats use only the ego- and allocentric strategies respectively in an elevated plus maze task (Packard and McGaugh, 1996). This well-documented allocentric mapping is central to the importance of the hippocampus across cognitive functions; encoding flexible representations of space and relevant features is needed for subsequent recognition and planning.

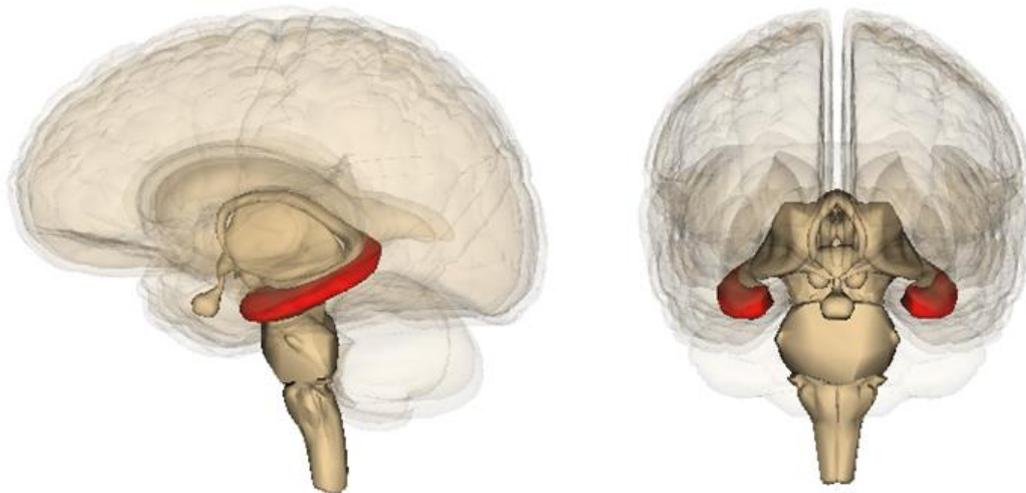
A**Allocentric****Egocentric****B**

Figure 1.7 Neural basis for spatial cognition.

A) Representations of allo- and ego-centric navigation perspectives. The hippocampus provides a system for building allocentric representations of space based on egocentric experiences. From <https://www.google.co.uk/maps>. **B)** Location and size of the human hippocampus (red). From Wikimedia Commons (September 2016).

Electrophysiology of the hippocampus

The hippocampus consists of two inter-locking magnetically 'open-field' structures: the dentate gyrus (DG) and Cornu Ammonis (CA consisting of subfields CA1-CA3). The pyramidal cells found in neocortex layer V and CA subfields of the hippocampus are morphologically indistinguishable (Spruston, 2008) (**Figure 1.8A**). In both pyramidal cell layers, the principal neuronal axes of the dendritic trees are arranged in parallel with one another, perpendicularly to the surface envelope. At a population level one

can therefore model current flow along the principal neuronal axis (red arrow in **Figure 1.7A**) in the same way as per convention for the neocortex. Although the hippocampal pyramidal cells point in the opposite direction to those in neocortex, this does not influence the shape or extent of the magnetic fields produced and therefore need not be explicitly modelled.

Conversely, dentate gyrus mainly consists of small granular cells which also have an oriented dendritic tree and a high cell density (Attal et al., 2007; Duvernoy, 2005). However, these cells are much smaller and shorter and are thus less likely to give rise to dipole-like sources. Moreover, it is difficult or impossible using 3T MRI images to segment these substructures accurately (Bonnici et al., 2012; Wisse et al., 2012). Other MEG models have made the simplifying assumption that the hippocampal signal originates only in the CA pyramidal neurons such that dipoles are modelled orthogonally to the surface envelope (**Figure 1.8B**, (Attal et al., 2012)).

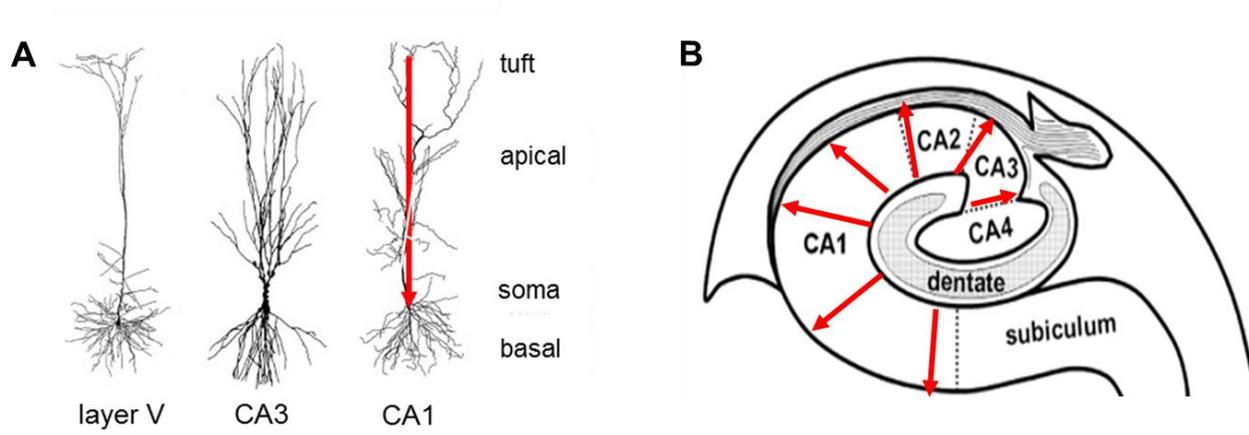


Figure 1.8 Hippocampal cell morphology and subfield structures

A) Morphology and similarity of pyramidal neurons in cortex and hippocampus. Postsynaptic potentials occurring at the apical dendrites or tuft give rise to the primary intracellular current (red arrow) which is measurable outside the head given a sufficiently large synchronously firing cell population. CA: Cornu Ammonis. Cells pictured are from the rat (but representative of all three cell types in humans). Image modified from (Spruston, 2008). **B)** Diagram showing distribution of subfields in a coronal plane. CA1-3 folds around the dentate gyrus and almost encapsulates it. Adapted from (Yang et al., 2008).

MEG source reconstruction in the case of the hippocampus

Cumulative evidence suggests that hippocampal sources can be identified in MEG, an observation made both with simulations (Attal and Schwartz, 2013; Chupin et al., 2002; Mills et al., 2012; Quraan et al., 2011; Stephen et al., 2005), and empirical data (Adjamian et al., 2004; Backus et al., 2016; Cornwell et al., 2012, 2008; Engels et al., 2016; Guitart-Masip et al., 2013; Hillebrand et al., 2016; Kaplan et al., 2012a; Korczyn et al., 2013; Mills et al., 2012; Moses et al., 2011; Poch et al., 2011; Quraan et al., 2011; Riggs et al., 2009; Tesche and Karhu, 2000). Despite this body of theoretical support and empirical evidence, the sufficiency of the spatial precision of MEG for deep source reconstruction is still being debated (Mikuni et al., 1997; Mills et al., 2012; Riggs et al., 2009) or assumed insufficient. The main reason is perhaps that although these authors claim to record signals from hippocampus, the ground truth is not available and so validation of these claims is difficult. Another reason is that arguments for hippocampal involvement typically rely on the spatial location of a statistical peak in traditional group level volumetric inference. Consequently, factors which have led such findings to be toned down from 'hippocampus' to 'medial temporal lobe' include image smoothness at this depth (Gross et al., 2003), intra-subject variability, head movement and in particular, co-registration error. Another argument against its detectability is that its cylindrical geometry could cause signal cancellation (Baumgartner et al., 2000; Mikuni et al., 1997; Stephen et al., 2005). However, it has been demonstrated that the cancellation is lower than expected even when sources on opposing subfields are simulated (Stephen et al., 2005). Perhaps most importantly, direct evidence comes from two studies showing that concurrent intracranial electrode recordings and MEG reveals that MEG sensors can reliably detect hippocampal theta oscillations (Crespo-García et al., 2016; Dalal et al., 2013a). These two studies form a critical piece of cross-modal evidence which, unlike the non-invasive neuroimaging studies, has ground truth available and shows that the MEG sensors detect hippocampal activity both using beamforming (Crespo-García et al., 2016), and at sensor-level (Dalal et al., 2013a). Another similar observation comes from separate iEEG and MEG studies where the same theta correlates have been observed invasively and non-invasively in response to the same VR (Bush et al., 2015; Kaplan et al., 2012a).

Another commonly used argument is that the hippocampus is simply too deep to produce a measurable signal. Estimates suggest that the distance between the centroid of the hippocampal mesh and the nearest sensor is ~8 cm (Chapter 2). Although this is deep relative to neocortical structures which are only a few centimetres from the MEG sensors, it is more superficial than other structures successfully imaged

using MEG such as the thalamus and brainstem (Attal and Schwartz, 2013; Coffey et al., 2016; Papadelis et al., 2012; Parkkonen et al., 2009; Wibral et al., 2013). Moreover, recent evidence suggests that the cell density, and consequently also current source density in the hippocampal pyramidal cell layer is at least twice that of the neocortex, which might compensate to some degree for its distance to the sensors (Attal et al., 2012; Murakami and Okada, 2015, 2006). Thus, it seems likely that a measurable signal is produced by the hippocampus (Crespo-García et al., 2016; Dalal et al., 2013a). The question then is how best to induce theta band oscillations/activity in the human hippocampus in order to test this hypothesis?

Since the late 1990s, virtual reality (VR) has enabled a broader and more ecologically valid approach to studying spatial navigation in humans (Maguire et al., 1999). Several important features make this technology ideally suited for human neuroimaging experiments designed to engage hippocampus. First, VR simulates real world navigation in a highly naturalistic way. Unlike table-top tasks where the subject is presented with allocentric representations from the beginning, VR makes it possible for subjects to build these representations based on egocentric information, matching real-world processes. Second, VR makes it possible to study the dynamic processing involved in real navigation processes such as planning, path integration, wayfinding etc. Third, VR environments have enabled experimenters to directly test the effects of manipulating environmental layouts and content (e.g. landmarks, novel objects) on navigation performance and strategies (Maguire et al., 1999).

Empirical validation of the ecological validity comes from research showing that cognitive maps built through VR closely resemble those acquired naturally: with practice, people learn to navigate inside a simulated building with equal level of performance as in the real building. Interestingly and in accordance with the requirement for naturalistic environments/surroundings, landmarks were found to improve performance (though a form of spatial anchoring or improved reference capacity), while abstract coloured pattern cubes were not (Ruddle et al., 1997). Nonetheless, VR has drawbacks such as limited field of view, lack of vestibular and proprioceptive feedback/engagement, movement execution through a keypad and potential software issues such as screen resolution and speed are worth taking into consideration for data interpretation (Maguire et al., 1999). Thus, while the ideal freely moving experimental set-up is not currently possible, VR provides a highly useful simulated version thereof (Shine et al., 2016).

Summary

In summary, the hippocampus is thought (among other things) to enable a viewpoint independent representation of physical space through computations dependent on the theta rhythm. These oscillations are very well-studied in rodents and possess a range of quantifiable relationships with running speed, novelty, synaptic plasticity, and the activity of place and grid cells. The question therefore is whether and how these theories can be translated to, and expanded on, in human neuroscience.

At present, these oscillations are difficult to study non-invasively in humans (using MEG). The two primary reasons for this are distance from the MEG sensors (around 8 cm) and complexity or difference of the neural architecture from that of neocortex. However, the depth is a solvable SNR problem, and the hippocampal pyramidal cell layer is almost identical to neocortical layer V. Given appropriate/optimised acquisition and analysis methods, it should therefore be possible to devise an objective test for hippocampal theta.

There is now strong evidence showing that hippocampal theta oscillations can be detected in MEG from epileptic patients by recording simultaneous MEG and intracranial hippocampal activity (Crespo-García et al., 2016; Dalal et al., 2013a). If it is possible to create use MEG as a reliable and non-invasive methods for studying hippocampal oscillations in humans, then the usefulness and relevance of this approach extends across cognitive, clinical and computational neuroscience questions.

Overall Summary

Current dipoles generated by groups of synchronously active parallel pyramidal cells produce instantaneous magnetic fields at the scalp. This enables us to sample cognitively relevant brain oscillations directly but non-invasively.

Magnetic signals generated by the brain are extremely small, which means that it is necessary to both shield the external magnetic field and use highly sensitive devices for detection. Magnetically shielded rooms provide the required passive shielding, while SQUID and OPM sensing technology are both extremely sensitive, and have additional hardware and software based methods for noise minimization.

Sources of neural activity measured by these sensors can be reconstructed by using a set of mathematical models to “invert” the data from the sensor to the source level. Various assumptions or hypotheses about the spatial distribution of these signals can be embedded in different algorithms and generative models of how the brain gave rise to the measured signals.

The human hippocampus contains a layer of pyramidal cells which strongly resemble cortical pyramidal cells and therefore are likely to give rise to similar signals. There are strong predictions about the temporal dynamics of these signals, specifically theta oscillations, which could be tested in MEG if the spatial resolution of the data could be improved.

Thus, the questions which remain unanswered and which I aim to address in this thesis are:

- 1) Does including a model of the hippocampus in the generative model help to explain hippocampal data?
- 2) In a probabilistic or Bayesian framework, how does this model perform relative to a null hypothesis model?
- 3) How specific is this advantage anatomically? For example, do cortical sources give false positive results? Do medial temporal lobe sources? How sensitive is the advantage to translations and rotations of the mesh?
- 4) Can we reliably use performance metrics to evaluate the model’s performance which are not limited to simulated data (i.e. which do not rely on ground truth)? If such a metric is appropriate for evaluating the goodness of fit, it can also be applied to empirical data and provide directly comparable results.
- 5) How do different newly developed inversion algorithms compare to more classical methods in this context, and which would be more appropriate for empirical analysis?
- 6) How does this model perform in the face of realistic empirical perturbations such as noise and/or co-registration error? How can we optimise our acquisition protocol to meet such requirements? Does this enable us to empirically measure hippocampal signals?

- 7) Can we expect further improvements from using better MEG sensors in the future? What are some of the potential practical pitfalls of this new technology? How can these be addressed?

Chapter 2

Experiment 1: Using generative models to make probabilistic statements about hippocampal engagement in MEG

Precis

The current general consensus (as outlined in the introduction) is that MEG cannot reliably be used to localize activity from deep sources such as the hippocampus. If this is the case, then we wondered whether an improvement to the anatomical modelling, such as adding a nested hippocampal manifold to the cortical mesh, would give way for an improvement in the model fit and therefore reveal itself as source-specific mesh preferences in a Bayesian evaluation framework.

Here we evaluate the differences between model fits with and without this hippocampal manifold included in the model. We do this across different inversion algorithms and report their properties.

Introduction

The aim of this part of the PhD was to develop a method to infer not where an activation peak appears to be, but rather *whether* a model which includes the hippocampus does a significantly better job than a hippocampus-free model (i.e., a “null” model) at explaining hippocampal activity. We address this question by comparing two generative models, both including the cortex but one with and one without the hippocampus also included. A generative model is an account of the putative origins of the signal. The models therefore enable formulation of competing hypotheses, and direct comparison hereof. This work echoes previous papers on the suitability of fMRI priors (Henson et al., 2009) and distinction between cortical laminae (Troebinger et al., 2014a), where for a given dataset we evaluate the evidence for two competing generative models which differ with respect to their anatomy. In this simulation study, we focus on explaining the method and testing its performance under different empirical constraints. We know from previous work that mesh-based generative models are extremely sensitive to co-registration error (errors in aligning an anatomical MRI image used to constrain the inverse solution, and the recorded MEG data) (Hillebrand and Barnes, 2011, 2003; López et al., 2012; Troebinger et al., 2014b) which therefore constituted our main factor of interest.

Here we propose an anatomically and electrophysiologically realistic generative model of deep source activity which accounts for geometry, depth and cell type. Through model comparison, this allows us to make categorical statements about which generative model is most likely for a given dataset – one with the hippocampus explicitly modelled, or one without. Although we focus on the hippocampus in this

work, the approach should generalize to other structures with similar features. Here the modelling is motivated by the similarities between the pyramidal cell layer V in neocortex which is the main generator of the MEG signal (Murakami and Okada, 2006), and the pyramidal cell layer of the hippocampus. Firstly, the cells are morphologically identical (**Figure 1.8A**). Secondly, the pyramidal cell layer follows the surface curvature which means that it can be modelled as such. Thirdly, individual cells have dendritic trees oriented in parallel, thus causing magnetic fields to arise perpendicularly to the surface.

The main advantage of an explicit generative model is that it makes it possible to exploit not only the information from the estimated source location but also its orientation (and other parameters not considered here like current density and local coherence). We will show that this allows us to differentiate the hippocampus from even the most proximal cortical sources.

In order to obtain probabilistic and comparative estimates of how good the two generative models are with respect to the data, we approximate their model evidence and compare the relative values in a Bayesian framework. This Bayesian model comparison uses these model evidence values and is a useful way to compare models because it allows direct quantification of competing models' abilities to explain the same data while avoiding over-fitting. Building models equates to specifying prior beliefs about what could be expected from the data. In this case, the priors pertain to the anatomical locations and orientations of the potential sources, and functional properties of sources, e.g. how sparse or smooth they are (different functional priors or inversion schemes).

To approximate the model evidence, we use Free energy (F), a lower bound on the true model evidence. F rewards models which accurately fit the data, but penalizes models based on their complexity. The former helps identify good hypotheses, while the latter eliminates over-fitting noise. The logic in this context is that if electrical current was generated on the hippocampus but the hippocampus is not part of the generative model used to reconstruct the data, then a more extensive mixture of cortical sources is required to explain the data equally well. Because of the increased complexity (see Wipf and Nagarajan 2009 on how the volume of the model covariance acts as penalty or sparsifying term), the hippocampus-free model will have lower model evidence (or Free energy) than model which includes the hippocampus.

The aim of this chapter is to first introduce the notion of a generative model, and then relate this to the simulation and source reconstruction procedures and parameters.

Next the model comparison framework is introduced, where two models are assessed in relation to each other, in a set of scenarios with different sources of uncertainty, specifically co-registration error and white noise added to the sensors. This model comparison is done across three different sets of popular functional priors or beliefs about the structure of the neural activity (e.g. how smooth and sparse it might be): Minimum Norm Estimate (MNE), Empirical Bayes Beamformer (EBB), and Multiple Sparse Priors (MSP). This allows us to interrogate the model comparison framework from multiple angles, as there is no single superior functional prior since the performance depends both on the experimental question(s), performance criteria, and data (Hauk et al., 2011). Moreover, this allows us to address the consistency of results across functional assumptions, i.e. the robustness to both different types of uncertainty, and different assumptions about source covariance.

Hypothesis and objectives

Hypothesis: If the hippocampus is explicitly incorporated into the generative MEG source model, then it is possible to test whether or not it is active at a certain time and within a certain frequency band by using Bayesian model comparison.

This relies on the validity of the Occam's razor approach: we assume that the simplest way of modelling a source is the correct one. Specifically, we show that if hippocampal activity is simulated, then a generative model which includes the hippocampus gives a more parsimonious, and therefore better, inverse solution.

Our objective here is to test the limitations and robustness of this approach in simulations where ground truth is known. A related objective is to quantify the effects of different empirical acquisition factors, namely co-registration error and SNR, on our ability to successfully and reliably detect hippocampal sources.

Methods

Anatomical modelling of the hippocampus

The independent variable of our generative model is the hippocampal surface mesh. We constrain the sources to be oriented perpendicularly to the mesh surface (**Figure 2.1A** shows the surface envelope extracted from an MRI image). The hippocampus' location is overlaid on an MRI image (**Figure 2.1B**) and shown with respect to the

cortical mesh (**Figure 2.1C**). As the hippocampus bulges into the floor of the (inferior horn of the) lateral ventricle, its medial surface extends more medially than that of the cortical surface. Apart from this, the hippocampus is nested inside the cortical manifold.

We extracted the left hemisphere's cortical and hippocampal surfaces for a single subject using FreeSurfer's (Reuter et al., 2012) automated image segmentation of individual T1-weighted MRI images (3T Siemens Magnetom). FreeSurfer gave a cortical mesh that we used directly, and a hippocampal volume file which we converted into a tessellated surface mesh. We limited the simulations and re-constructions to the left side of the brain for simplicity. The resultant hippocampal surface was more densely tessellated than the cortical, so we smoothed and downsampled it such that the mean vertex-vertex distances matched. The number of vertices in the cortical and hippocampal meshes were 10595 and 162 respectively and the mean vertex-vertex distances were 3.73 and 3.69 mm. This approach is consistent with the Deep Brain Activity model proposed by (Attal and Schwartz, 2013).

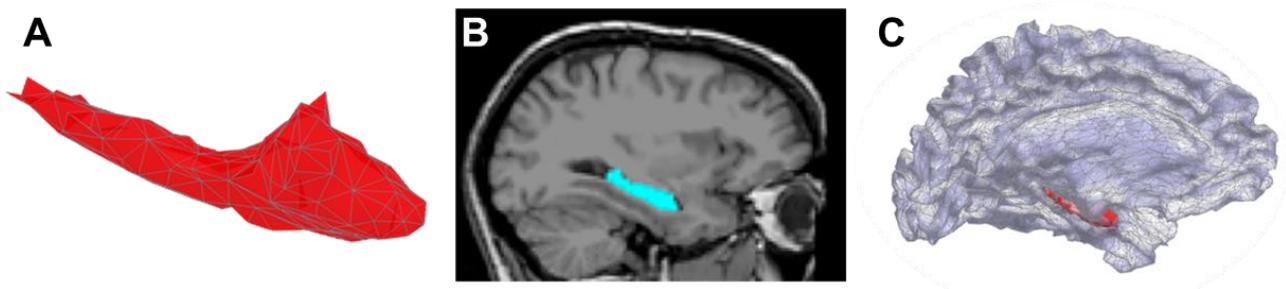


Figure 2.1: Hippocampal surface structure and location

A FreeSurfer-derived tessellated envelope of the left hippocampus. We model the sources to be perpendicular to mesh vertices, consistent with the pyramidal cell orientation. **B** Sagittal view of FreeSurfer hippocampal region of interest on a sample 1.5T T1-weighted MR image from the FreeSurfer Image Analysis Suite. Blue colour shows the extent of hippocampal region of interest. Image adapted from (Hostage et al., 2013) **C** Source space of the combined model consisting of FreeSurfer-derived cortical and hippocampal meshes. For more detailed description of this model, see Figure 3.

Simulation set-up

The simulation and reconstruction pipeline consisted of three steps: first, we simulated a single dipole perpendicularly to the hippocampal surface with a sinusoidal waveform of 20 Hz for 300 ms (six cycles) and a total effective dipole moment of 20 nAm (**Figure 2.2A**). The simulation locations were randomly drawn from the 162 hippocampal vertices and were simulated with a full-width half-maximum of 6 mm. Each simulated dataset had a sampling rate of 600 Hz with the mean sensor-level Signal to Noise Ratio (SNR) set to either 0, -5, -10, -15 or -20 dB, specified by adding Gaussian white noise to the data. We repeated this 30 times with both hippocampal and cortical simulation locations at each SNR level. This gave a core set of simulated data with known ground truth (hippocampal source or not).

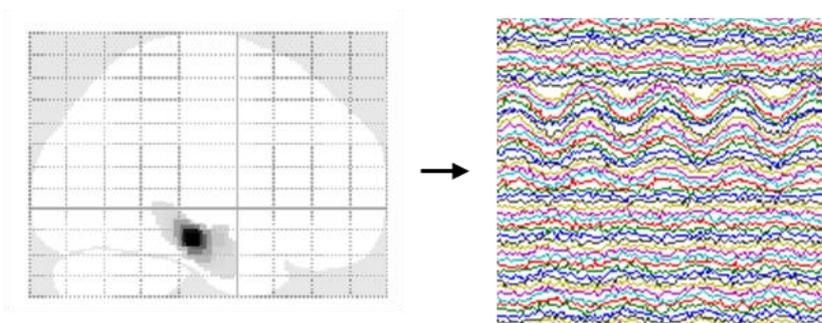
In the second step (**Figure 2.2B**), we mimicked the effect of co-registration error between functional (MEG) and anatomical (MRI) images by adding 0, 1, 2 or 3 mm standard deviations of error to each of three fiducial points in each of the three spatial dimensions. This shifted the surface mesh used for reconstruction (red) relative to the surface mesh used to generate the simulation (black). Co-registration error levels commonly seen empirically in MEG recordings are usually ~5 mm or more even with the best compensation tools, be they bite-bars (Adjamian et al., 2004; Singh et al., 1997) or algorithmic movement corrections (Whalen et al., 2008).

After having perturbed the idealized data by adding sensor noise and co-registration error, we inverted the data using two different anatomical models and three different inversion schemes. One anatomical model was, per convention, just the cortical surface (**Figure 2.2C**, cortical model), while the other model additionally included the hippocampal surface envelope (**Figure 2.2C**, combined model). Each anatomical model was inverted using three different inversion schemes embodying functional (or source covariance) assumptions. These were Minimum Norm Estimate (MNE) (Hämäläinen et al., 1993), Empirical Bayesian Beamforming (EBB) (Belardinelli et al., 2012) and Multiple Sparse Priors (MSP) (Friston et al., 2008a). We thus obtained six inversion solutions per simulated dataset; three inversion algorithms, each giving one solution per anatomical model.

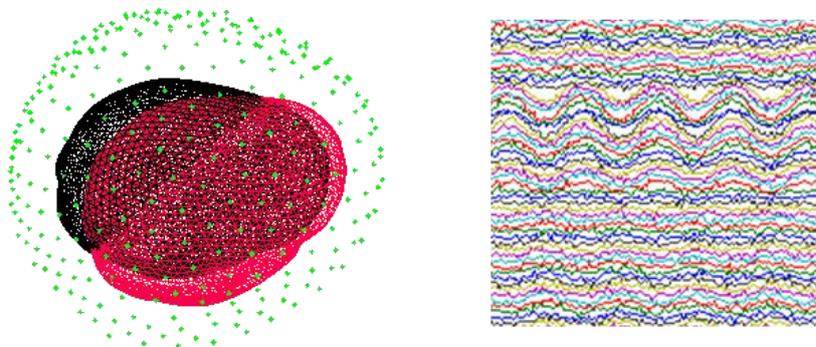
This lets us examine the difference between generative models across different assumptions about the nature of the activity – how sparse, how co-varying, how smooth etc. Each such inversion returns a Free energy (F) value, which approximates the model evidence for generative model. This set-up allowed us to quantify the difference in model evidence when the hippocampal mesh is included in the generative model. The hypothesis was that there would be an improvement if the

simulated source was hippocampal. This model comparison approach has successfully been demonstrated elsewhere (Henson et al., 2011, 2009; Lopez et al., 2013; López et al., 2014; Penny, 2012; Stevenson et al., 2014; Troebinger et al., 2014a). Here we used log Free Energy to quantify the difference between anatomical models: $\Delta F_{\text{anatomical}} = F_{\text{combined}} - F_{\text{cortical}}$. A positive difference means that the combined model is $\frac{1}{1+e^{-\Delta F}}$ more likely than the cortical. If $\Delta F = 0$ then the two models are equally likely, and if $\Delta F = 3$ then the combined model is approximately twenty times more likely.

A Simulate single hippocampal source



B Add co-registration error without changing data



C Invert data with two anatomical models

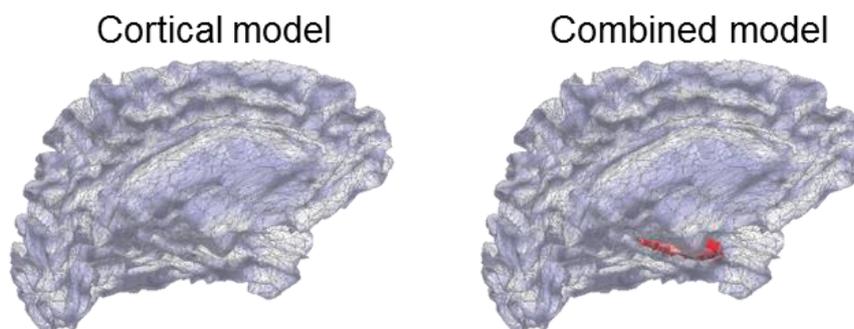


Figure 2.2: Overview of the simulation pipeline.

A A single dipole source is simulated (at a random location) on the hippocampal surface as a temporal waveform with sinusoidal frequency of 20 Hz. Gaussian white noise is added to the sensor level data (in this case -10 dB). On the right, a representative subset of the resulting 274 time-varying waveforms simulated are shown as coloured traces. **B** To simulate the effects of co-registration error, we added a displacement of 0, 1, 2, or 3 mm standard deviation of error in each spatial dimension to each of the three standard fiducial points. The data themselves were unchanged. The displacement shown here is 2 cm for illustration. **C** Next we inverted the simulated data twice, using two different generative models. One with only the cortical surface (cortical model) and one with both cortical and hippocampal surfaces (combined model). We repeated this double inversion procedure on each dataset using three different reconstruction algorithms.

Specification of anatomical priors

The schematic in **Figure 2.3** illustrates the two anatomical models and how they were implemented. The key difference is that MSP priors can be user-defined within subsections of the source space. Conversely, EBB and MNE by definition make use of the complete source space. Left panels (A and C) show the cortical models and right panels (B and D) show combined models (with hippocampal priors). For EBB and MNE, the addition of hippocampal priors simply involves an addition to the source space (which increases from 10595 vertices to 10757 vertices). For MSP on the other hand, we kept the complete source space (combined model with 10757 vertices) but specified 100 spatial priors (patches of cortex) to either include or not include the hippocampus. The 90 blue dots mark cortical priors shared across the two models. The ten green dots mark cortical priors unique to the cortical model (**Figure 2.3C**). The ten red dots mark hippocampal priors unique to the combined model (**Figure 2.3D**). In all cases, we used a Nolte single shell (Nolte, 2003) to model the inner skull boundary.

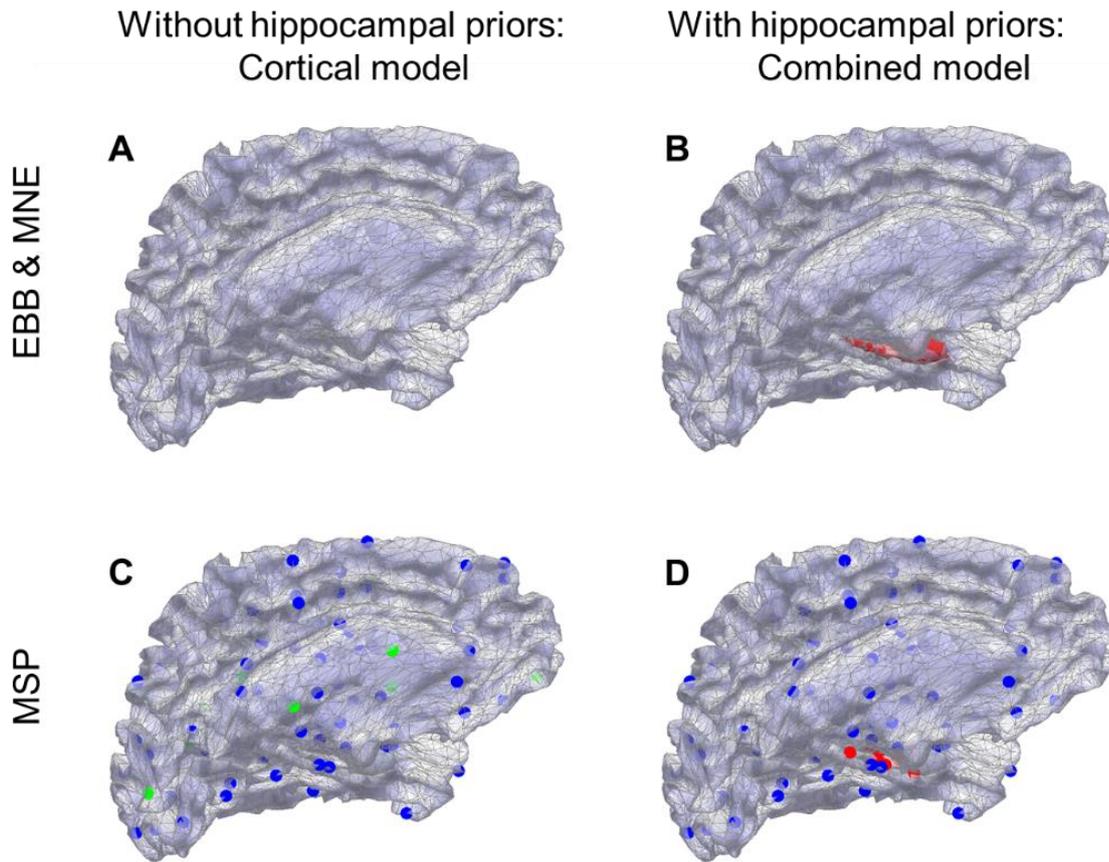


Figure 2.3: Anatomical models with and without hippocampal priors.

Panel **A** shows implementation of the cortical model in the EBB and MNE algorithms. The tessellated cortical surface envelope is comprised of 10595 vertices. Panel **B** shows the combined model which includes a nested hippocampal manifold and contains 10757 vertices. Bottom panels (**C**, **D**) show the anatomical model implementations in MSP. The full source space is specified in both models such that each includes the nested hippocampal mesh and the number of vertices is 10757. In both, 90 blue dots illustrate identical cortical prior locations. In panel **C** an additional ten green cortical priors are specified. In **D**, an additional ten red hippocampal priors are specified.

Source Inversion

The empirical Bayes source inversion scheme has been described in detail elsewhere (Belardinelli et al., 2012; Friston et al., 2007; Henson et al., 2011; López et al., 2012; Phillips et al., 2005; Troebinger et al., 2014a). For a review, see (López et al., 2014). Here we elaborate on implementation issues and empirical applications.

All three algorithms require the estimation of a source and sensor level covariance matrix. In all cases we used an identity matrix to represent uncorrelated white sensor level covariance. The main difference between the three algorithms is that the MNE and EBB solutions require the optimization of a single source level covariance prior whereas MSP has a more general form. In MNE this is also an identity matrix (all sources have equal prior variance and are uncorrelated); for the EBB algorithm this prior is derived directly from the data. This means that for EBB and MNE the algorithm must estimate two (hyper) parameters which set the relative weighting of source and sensor level variances. The MSP algorithm takes a more general form and allows the source distribution to be built up of multiple covariance components. Traditionally each of these components is a locally coherent patch of cortical activity. The ensuing optimisation (to maximize Free energy) can be thought of as a process to minimize number of patches but still explain the maximum amount of data. The mixing and pruning of these priors means that for large numbers of priors, the optimisation can get trapped in local extrema.

One practical solution to avoid this is to run the same algorithm many times with different sets of priors (Troebinger et al., 2014a). However, as we were not interested in the optimisation per-se in this work, but in finding the best possible solution, we used 100 priors and simulated sources at a subset of these locations. Note that there was thus a clear advantage for the MSP algorithm relative to EBB and MNE, because the best solution is fixed to lie in the space of MSP priors, which is much smaller than the space of all the vertices (See **Figure 2.3** and discussion). This advantage is relevant in both hippocampal and cortical simulation results. For hippocampus, the 10 MSP priors included the simulated patch, versus all 162 hippocampal vertices supplied with EBB/MNE. Similarly for the cortical simulations, the solution space was defined by 90 anatomical priors for MSP, again including the simulated patch, versus all 10595 cortical vertices specified for the EBB/MNE algorithms. Importantly, it is still possible to directly compare the inversion schemes by keeping the model (and data) constant. We return to this and examine it across a range of SNRs.

We did not use any spatial dimension reduction (i.e. all 274 functioning MEG channels were used) but we decomposed the time series into a single temporal mode. The time window was set to match that of simulation (0-300 ms), as was the frequency band of interest (0-80 Hz, simulated waveform of 20 Hz). A Hanning taper was applied to the time series. We used three different forms of functional priors (MNE, EBB and MSP) and two sets of anatomical priors (cortical and combined models). Sample inverse

solutions for all six prior combinations are shown in **Figure 2.4A**. We carried out 30 iterations of each hippocampal and cortical simulations at each SNR level.

Dipole Localisation Error Analysis

In order to provide a frame of reference between the model evidence based approach and other simulation studies we also calculated the dipole localisation error (DLE). The DLE equates to the distance between the true simulation location and source distribution maximum of the inversion. The latter was defined as the peak in the estimated primary current density matrix. We calculated DLEs separately for the combined and the cortical models used to invert 30 hippocampal and (random) cortical simulation scenarios using EBB.

Results

Variance Explained and Free Energy

In order to demonstrate the basic logic behind our analysis **Figure 2.4A** shows a representative single-simulation source reconstruction for each combination of anatomical and functional priors. We can compare the algorithms qualitatively with respect to accuracy and complexity because we know the true source location. Spatial accuracy can be assessed by looking at how far the simulation vertex (red circle) is from the peak (darkest vertex) of the estimated current distribution. The complexity is reflected in the spread of the source estimates. Note that when the correct anatomical model is used (**Figure 2.4A**, top row), for EBB and MSP, the source estimates are generally accurate and focal. The increase in complexity (most noticeable for MSP and EBB) in the bottom row (inversions using just the cortical model) occurs because it takes more non-hippocampal sources to describe MEG data arising from a single hippocampal source. The simulation used here has sensor-level SNR -5 dB and zero co-registration error added.

We find that as expected, MMN gives the most diffuse solution and MSP and EBB give the most focal. Nonetheless, it is encouraging to note that although the algorithms have different functional assumptions, the estimated activity is in approximately the same place throughout.

In contrast to F , variance explained is not penalized for complexity and consequently is not discriminative of the correct model. **Figure 2.4B** illustrates the mean percentage of variance explained over 30 iterations of hippocampal simulations with SNR -5 dB whereas **Figure 2.4C** illustrates the mean Free energy. Note that the mean variance explained is >99.5% for all algorithms, and that the best model in terms of Free energy (MSP) does not explain the most variance. This is because there is less over-fitting of the noise.

Given that the Free Energy values do not rely on information about the true source location, it is ideally suited for evaluation of empirical as well as simulated data. For example, it has been shown that Free energy correlates with cross-validation accuracy as demonstrated by machine learning evaluations (Penny and Roberts, 1999), and with conventional reconstruction evaluation measures such as dipole localization error (Belardinelli et al., 2012). Although we *do* have access to the ground truth in these simulations, we will nonetheless rely on Free energy as a goodness of fit criterion but also evaluate the dipole localisation error for comparison. The main focus will be evaluation of two forms of Free energy differences, shown in **Figure 2.4C**. The bars encode mean Free energy values over 30 iterations of hippocampal simulations with SNR -5 dB. We first compare anatomical priors by subtracting the two Free energy values obtained using different anatomical models with the same algorithm. This is shown for MSP where $\Delta F_{\text{anatomical}} = F_{\text{combined}} - F_{\text{cortical}}$. We then compare functional priors by subtracting the two Free energy values obtained using the same anatomical model but different algorithms. For example, comparing Free energy with the combined model using EBB and MSP: $\Delta F_{\text{functional}} \text{ MSP vs EBB} = F_{\text{MSP}} - F_{\text{EBB}}$. This metric tells us how good the functional assumptions are (how smooth/sparse etc.), because the data and anatomical model are constant (the results of these tests are shown in **Figure 2.11**).

The main emphasis of this paper is on $\Delta F_{\text{anatomical}}$, or quantifying hippocampal engagement probabilistically through comparison of generative models. With respect to single-simulation $\Delta F_{\text{anatomical}}$ values corresponding to solutions shown in **Figure 2.4A**, we find that for all three algorithms, the combined (true) model has a higher F than the cortical model (single simulation $\Delta F_{\text{anatomical}}$ MMN = 1.4, EBB = 10.6, MSP = 73.2). We find that the average $\Delta F_{\text{anatomical}}$ values across 30 simulations (**Figure 2.4C**), are somewhat similar (mean $\Delta F_{\text{anatomical}}$ MMN = 1.0, EBB = 6.0, MSP = 23.1). Note that only EBB and MSP pass the significance threshold of 3 (log units). Thus, even without knowledge about true simulated source locations, Bayesian model comparison can distinguish between anatomical models, and thereby be used to infer whether the

source location is hippocampal or not. Interestingly, in this example EBB appears (from the source level maps) to perform equally well for both anatomical models. One explanation for why the peak of the cortical model solution appears to be in/on the hippocampus when it is not explicitly modelled (**Figure 2.4A**), is that the cortical and hippocampal mesh surfaces are very close together (see **Figure 2.3B**). Since EBB can distribute variance across all source vertices, those on the medial temporal lobe could therefore appear hippocampal. This issue is directly addressed later in **Figure 2.8**. Note that the performance of algorithms with certain models is a separate question from $\Delta F_{\text{anatomical}}$,

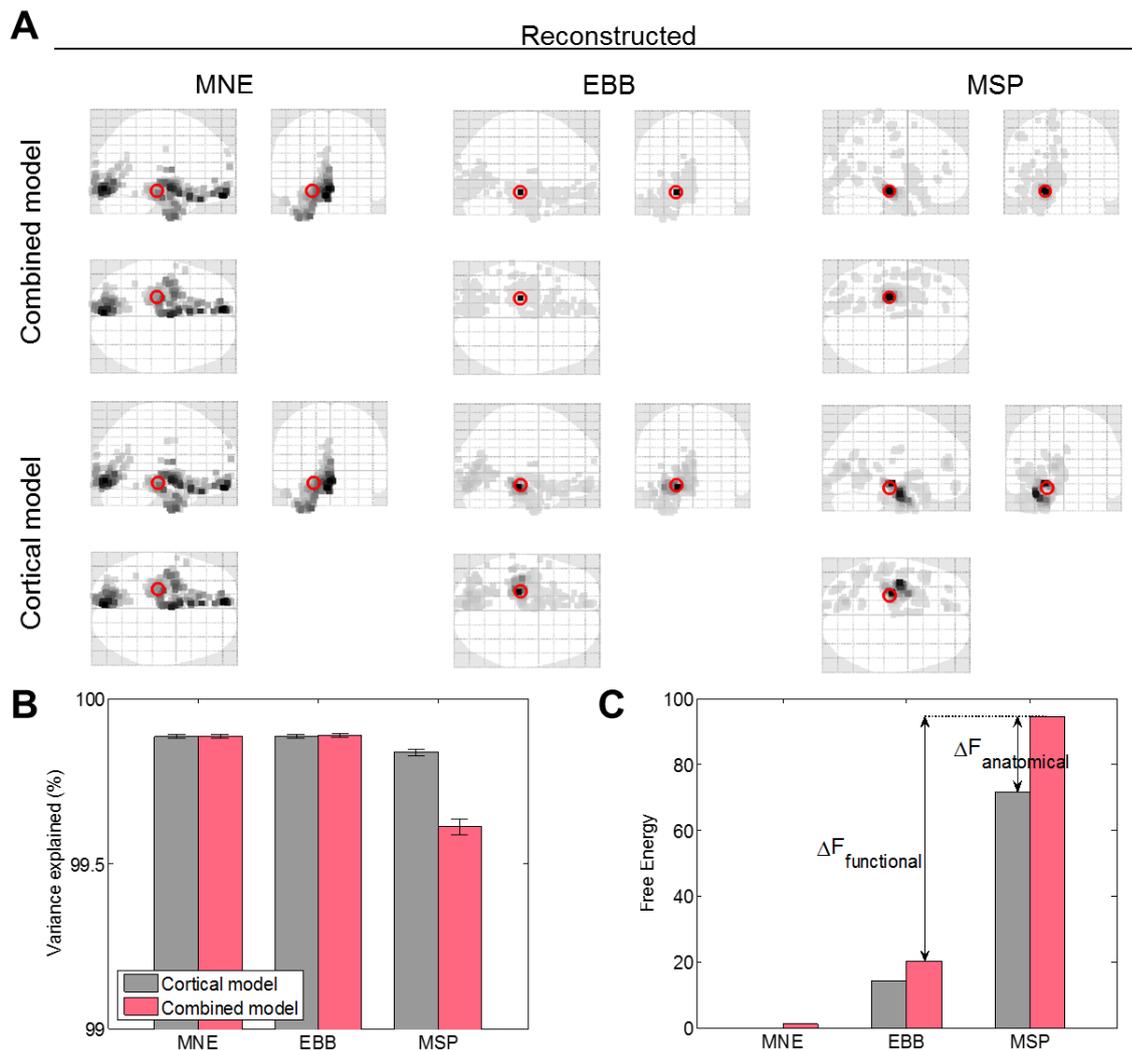


Figure 2.4: Sample source reconstructions and model comparison.

A) Single-trial reconstructions of a hippocampal source (red circles) with MNE, EBB and MSP priors using the combined model (top row) and the cortical model (bottom row). EBB and MSP accurately capture the true source location. Glass brains show estimated current source

density with the grey scale proportional to the darkest (maximally active) vertex location. Sample source simulated with SNR -5 dB and no co-registration error. **B)** Variance explained by different anatomical and functional priors. Bars encode mean percentage variance explained across 30 hippocampal simulations (\pm SEM). Note that the y axis only spans 99-100%. For this metric there was no significant differences between models with EBB ($t(29) = 1.0842$, $p = 0.287$) or MNE ($t(29) = 0.1591$, $p = 0.875$). For MSP in contrast, there was a significant difference in the percentage variance explained ($t(29) = -8.6310$, $p < 0.001$), but favouring the incorrect (cortical) model. **C)** Bayesian model comparison methods. Free energy (F) is used to approximate model evidence. Bars encode mean Free energy values over 30 simulations, normalized to MNE cortical. Differences between anatomical priors we denote $\Delta F_{\text{anatomical}}$ whereas differences arising from different functional priors we denote $\Delta F_{\text{functional}}$.

Anatomical Model Comparison

We evaluated two variations of the same basic generative model, one that included a nested hippocampal manifold and one that did not. To verify that the combined model helps to explain hippocampal activity, we simulated hippocampal sources and compared the Free energy values obtained with the two anatomical models ($\Delta F_{\text{anatomical}} = F_{\text{combined}} - F_{\text{cortical}}$). We observed that as expected, the combined model increased model evidence. **Figure 2.5A** shows the positive $\Delta F_{\text{anatomical}}$ values from across 30 simulated hippocampal datasets with SNR of -5 dB and zero co-registration error. As a first control, we tested whether this improvement was anatomically specific or could be driven by an increase in vertices regardless of the source location. We therefore simulated cortical sources and evaluated them in the same way before. The cortical sources were randomly distributed across the cortical mesh and again the simulation locations equated to (30 of the cortical) MSP priors. Given that the locations of the cortical priors (sparse or mesh-wide) were identical in the cortical and combined models, we expected to find no difference in model evidence between anatomical models. **Figure 2.5B** shows the null $\Delta F_{\text{anatomical}}$ values for data simulated on the cortical surfaces.

In order to derive a conservative bound on whether the models differed significantly, we computed the the Bayes Omnibus Risk (BOR) which quantifies the probability that the null hypothesis (that there is *not* a true difference in model frequency of winning) is true and any observed differences between models observed are due to chance (Rigoux et al., 2014). Table 1 shows the mean Free energy differences and accompanying BOR values across all algorithms tested for the hippocampal and

cortical control simulations. The mean Free energy difference must be above three while the BOR must be below 0.05 to reach significance (i.e. conclude that the results are only 5% likely to have been obtained by chance). While MSP and EBB both have a mean Free energy difference above 3 and a significant BOR value (and thus show significantly improved model fits with the combined model), the MNE Free energy mean does not surpass significance although the BOR is significant. For the cortical simulations, none of the mean nor BOR values reach significance.

Table 1: Bayes Omnibus Risk values for hippocampal and cortical simulations

	Hippocampal simulations		Cortical simulations	
	Free energy mean	BOR	Free energy mean	BOR
MSP	23.09	<0.001	-0.0490	0.8011
EBB	6.01	<0.001	0.0951	0.7930
MNE	1.04	<0.001	0.0642	0.7862

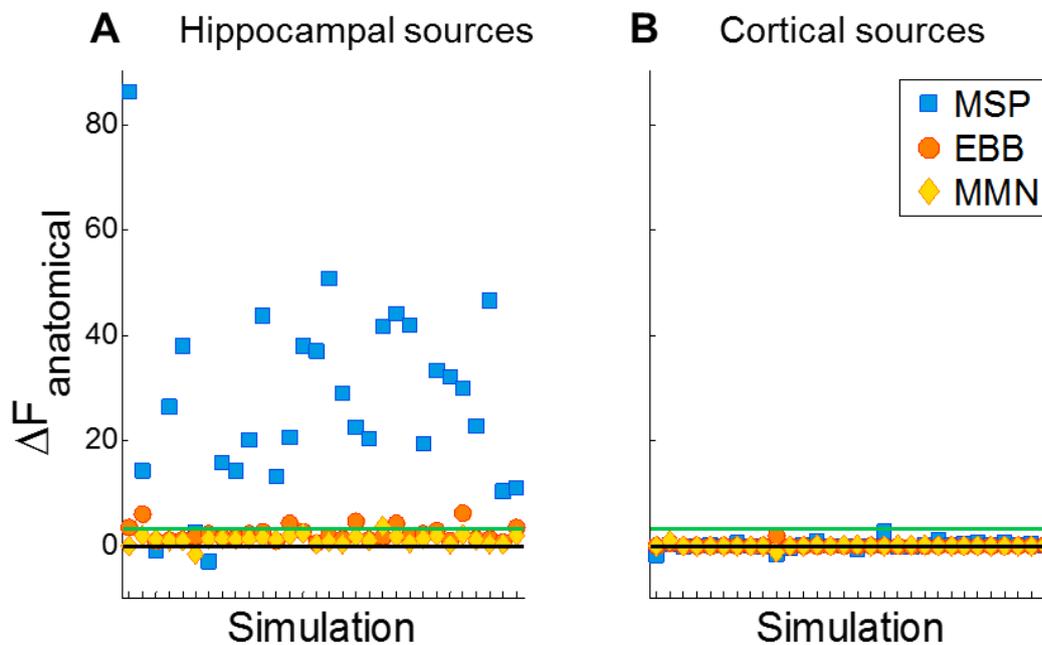


Figure 2.5: Anatomical model comparison for hippocampal and cortical (control) sources.

A Dots show $\Delta F_{\text{anatomical}} = F_{\text{combined}} - F_{\text{cortical}}$ values for sources simulated on the hippocampus. $\Delta F_{\text{anatomical}}$ is positive because the combined model explains more data using fewer (hippocampal) priors. The black line marks zero where there is no difference between models. The green line marks a positive difference of 3 which, because F is on a log scale, means that

the combined model is >20 times more likely than the cortical. MSP outperforms the other algorithms while MNE fails to reach significance. **B** Shows the results for the simulated cortical sources or control condition (note that here exactly the same comparison between full and cortical models is made). There is little if any difference between models because the models contain the same cortical mesh (all 10595 cortical vertices for EBB and MNE) or cortical priors (90/100 priors for MSP where the hippocampal priors are redundant and therefore pruned away in the combined model). For the 30 hippocampal and 30 cortical simulations shown, SNR is -5 dB and no co-registration error is added.

Effect of Co-registration Error

We then examined the effect co-registration error on our ability to identify the correct model. To do this, we simulated co-registration error by adding 0, 1, 2 or 3 mm standard deviation of error to each of three fiducial locations in each of three dimensions before inverting the model (see **Figure 2.2B**). Note that the shift and data were always the same for the two models. **Figure 2.6A-C** shows the model evidence differences obtained for the 30 hippocampal simulations described previously but with different levels of co-registration error. As expected, ΔF decreases as co-registration error increases, demonstrating that uncertainty about head location compromises our ability to evaluate and discriminate between models. We also found that the variability of ΔF values increases, illustrated most clearly with MNE (**Figure 2.6C**).

To quantify this we used a random effects analysis (Stephan et al., 2009) to estimate the probability that the correct (combined) model would win given a randomly drawn simulation run (grey lines, **Figure 2.6D-F**). Consistent with the model evidence difference decreases in the top panel, this probability decreases as co-registration error increases. If we were to select a dataset at random, we would expect to make the correct decision ~95% of the time with MSP, regardless of co-registration error. With the EBB this chance would decrease to ~75% at 3 mm of error and with MNE, we would be at chance level with 2 mm of error. One problem with this inference is that there is an underlying assumption that one model is better than another. In order to derive a (conservative) bound on where the models differed we again computed the the Bayes Omnibus Risk (BOR) which quantifies the probability that the null hypothesis is true and that differences between models observed occurred by chance (Rigoux et al., 2014). BOR probabilities (green lines in **Figure 2.6D-F**) of less than 0.05 (red lines) mean that the null hypothesis can be rejected. This showed that just 3 mm of co-registration error abolishes our ability to distinguish between models with

EBB and MNE. In sum, increased co-registration errors of ~3 mm or larger blur out existing differences between the anatomical models. Note that these co-registration errors are often observed and most often exceeded in conventional MEG recordings. Moreover, the closer the functional prior to the ground truth (compare MSP and MNE), the more robust it will be to co-registration error.

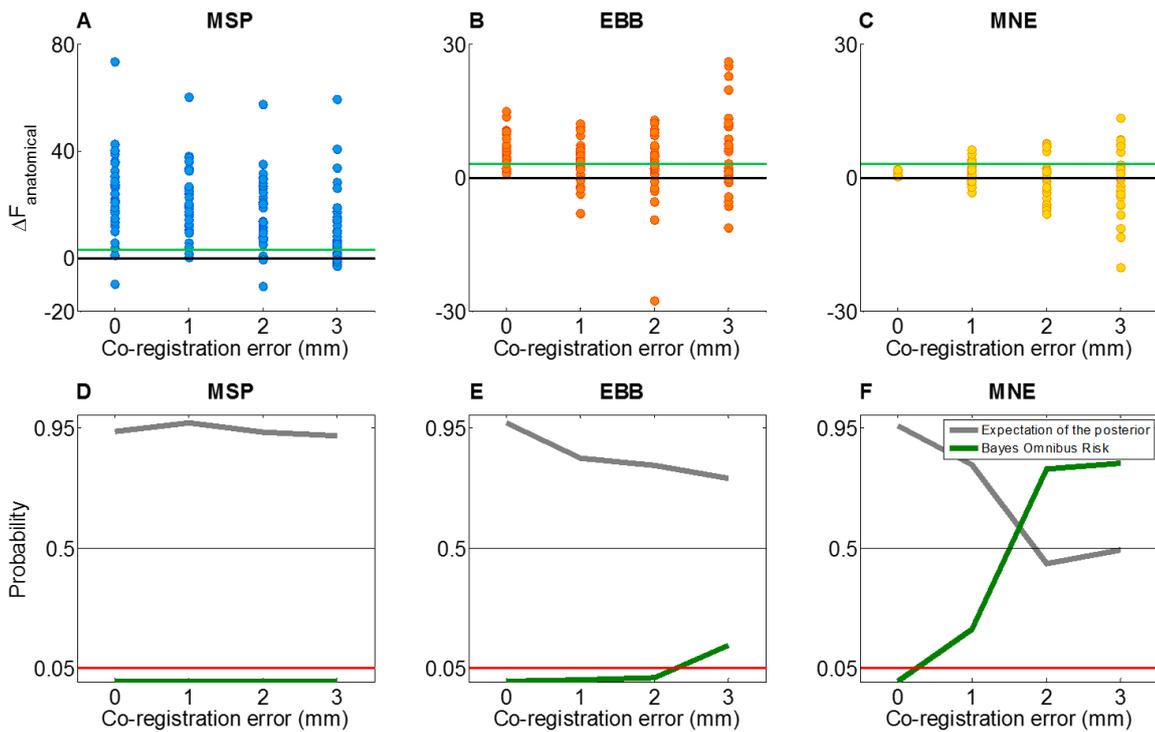


Figure 2.6: Effect of co-registration error on anatomical model comparison.

Inversion results from simulated hippocampal dipoles with SNR -5 dB and 0, 1, 2 or 3 mm standard deviation of error added to each of three fiducial locations in each dimension. **Top panel (A-C):** Dots represent $\Delta F_{\text{anatomical}}$ for the same 30 simulations at each co-registration error level. There is an increase in number of negative $\Delta F_{\text{anatomical}}$ values (false negatives) as a function of co-registration error. Green line marks the significance threshold of 3, black line marks no difference. Y-axes of EBB and MNE plots are adjusted for visibility. **Lower panel (D-F)** is structured in the same way but depicts two measures of the reliability of the model comparisons shown above. Grey line marks the expectation of the posterior; the probability that the combined model supersedes the cortical model. Green line marks the Bayes Omnibus Risk, the probability that anatomical model frequencies are equal (i.e. there is no difference between models); we can reject this null when this metric is below 0.05 (red line).

Effects of Co-registration Error and Sensor-level SNR

We next investigated the interaction between sensor level noise and co-registration error. We added different amounts of uncorrelated white noise to obtain 0, -5, -10, -15 and -20 dB SNR at sensor level. **Figure 2.7** takes the same form as **Figure 2.6** but includes an SNR dimension. The upper panel shows mean $\Delta F_{\text{anatomical}}$ over 30 hippocampal simulations where positive values show evidence in favor of the combined model. As expected, we find that as both co-registration error and noise increase, $\Delta F_{\text{anatomical}}$ decreases. The lower panel shows the Bayes Omnibus Risk quantified based on 30 hippocampal simulations at each combination of SNR and co-registration error. Green bar tops mark values BOR<0.05 where we can reject the null hypothesis that the models are equivalent, red bar tops mark the opposite (i.e. no difference between models). In general, we find that poor SNR is less detrimental to our ability to distinguish sources than co-registration error is. As before, we conclude that co-registration error must be <3 mm to make reliable identification of hippocampal activity with EBB and MNE. As expected (or defined by our simulations), the MSP outperforms the other two algorithms at all levels of co-registration error and SNR tested here.

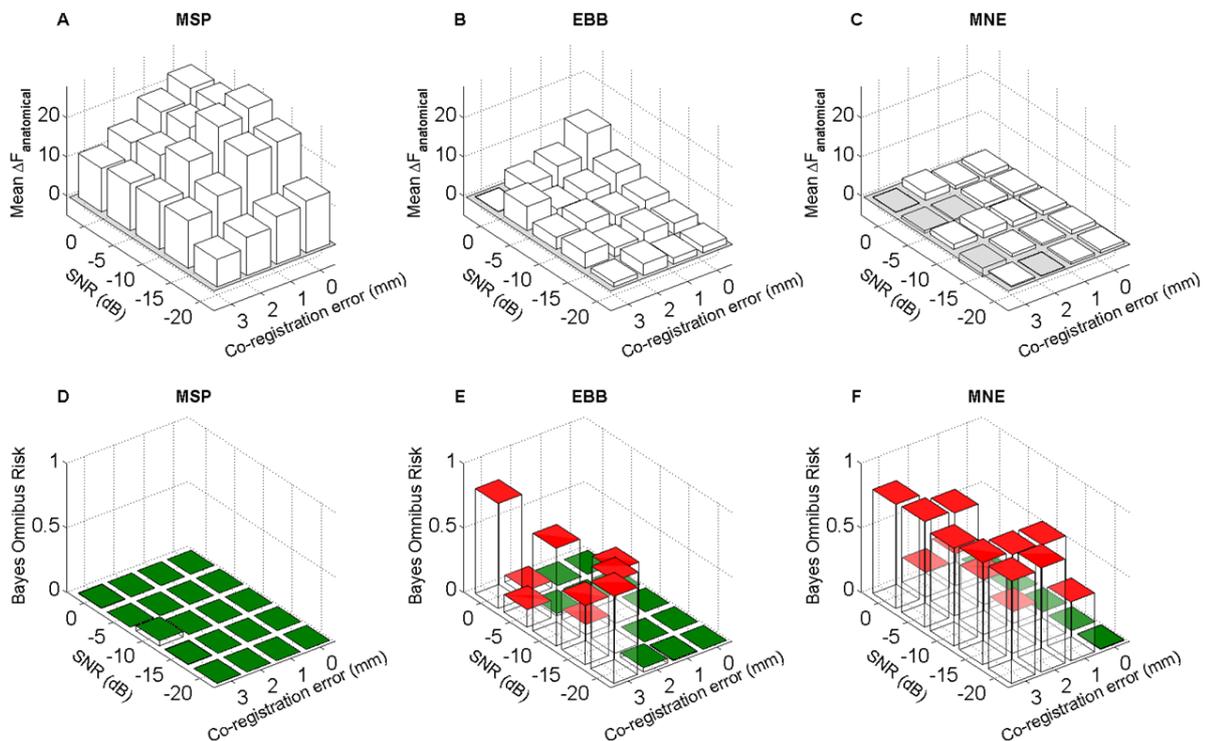


Figure 2.7: Effects of noise and co-registration error on anatomical model comparison.

The figure is similar to figure 6 with an added dimension of noise. **Top panel (A-C)** shows negative effects of co-registration error and noise: $\Delta F_{\text{anatomical}}$ decreases as a function of either. Each bar encodes average $\Delta F_{\text{anatomical}}$ of 30 reconstructed hippocampal simulations. **Lower panel (D-F)** shows roughly the same effects on the Bayesian Omnibus Risk, the risk that anatomical model frequencies are equal. Co-registration error above 0 and 1 mm are detrimental for MNE and EBB respectively. Green and red bar tops mark signify when the null (that there is no difference between models) can be rejected (BOR values <0.05) and not rejected respectively.

Closest Cortical Neighbours

As spatial resolution decreases rapidly with depth in MEG, there is a risk that higher Free energy values for the combined model could arise from nearby but non-hippocampal sources, yet be misinterpreted as hippocampal activity through the inference.

We tested this by simulating activity on the nearest cortical vertices to each of the 30 hippocampal vertices used in the original simulations and inverting these data with both the cortical and combined models to calculate model evidence difference for each location. Reassuringly, we found the average ΔF for the closest cortical neighbour simulations to be non-significant (mean 1.75, BOR <0.001) (**Figure 2.8A**, grey dots). Conversely, the hippocampal simulations gave positive and significant ($\Delta F > 3$, mean 6.01, BOR <0.001). The average distance between neighbouring hippocampal and cortical vertices was only 2.14 mm (**Figure 2.8B**).

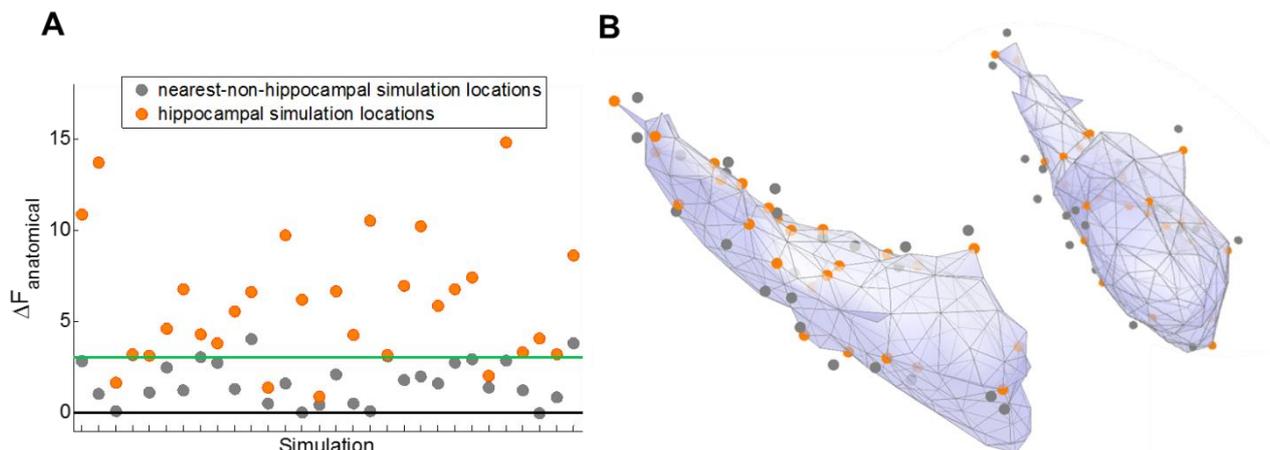


Figure 2.8: Closest cortical neighbour analysis.

A Orange dots reflect ΔF when activity is simulated on the hippocampal mesh (30 different sources shown here). Grey dots reflect ΔF when activity is simulated on the cortical surface but centred at the nearest cortical vertex to its hippocampal neighbour. Dots are vertically aligned in pairs (or neighbours). Simulating hippocampal simulations sources gives significant (>3 , green line) ΔF values whereas simulating on the nearest cortical neighbour does not. Parameters used were no co-registration error, SNR -5 dB and EBB. **B** Simulation locations visualised on two views the hippocampal mesh. Orange dots are on the hippocampal surface, grey are on the cortical surface (not visualised).

Effects of Translating the Hippocampus

To ensure that Free energy differences were specific to the correct model and not simply to having a deep structure added, we carried out a set of inversions with models that had the hippocampal slightly offset relative to the correct location. For this analysis, we used the same simulated hippocampal data as described previously (i.e., activity simulated on the hippocampal surface in its original location), but inverted these data using combined anatomical models with the hippocampal mesh slightly offset from the correct location (0.5, 1, 1.5 and 2 cm shifts) in three dimensions (medial-lateral, anterior-posterior, dorsal-ventral), and two directions (+ and -) giving 24 different shifted models (**Figure 2.9**). Note that the cortical portion of the combined model stayed the same. We focused here on EBB because its performance was mid-range and because it does not require specification of priors. We used simulations with SNR -5 dB. We inverted each of the 30 datasets with each of the 24 shifted models and compared the resulting Free energy values to those obtained with the standard cortical model as well as standard combined model. Only in cases where there is no translation (i.e. the correct combined model is used, middle bars), or there is 0.5 cm lateral translation, is the model comparison significant (all BOR values are significant, Table 2). This demonstrates specificity of the model comparison approach, and the ability to identify the correct model among a set of subtly offset alternative models. In other words, despite the physical overlap between surfaces when the hippocampus is translated, the disparity in the surface orientations mean that these shifted surfaces are poor generative models.

Table 2 shows the BOR values accompanying the translated mesh analysis. In all cases, the BOR is significant but only in the no shift and 0.5 cm lateral shift conditions is the mean Free energy difference between the two models greater than 3 (light green line, **Figure 2.9**).

Table 2: Bayes Omnibus Risk values for hippocampal translations

BOR	2 cm	1.5 cm	1 cm	0.5 cm	No shift	0.5 cm	1 cm	1.5 cm	2 cm
Medial-lateral	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	0.009
Anterior-posterior	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	0.003
Dorsal-ventral	0.012	0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001

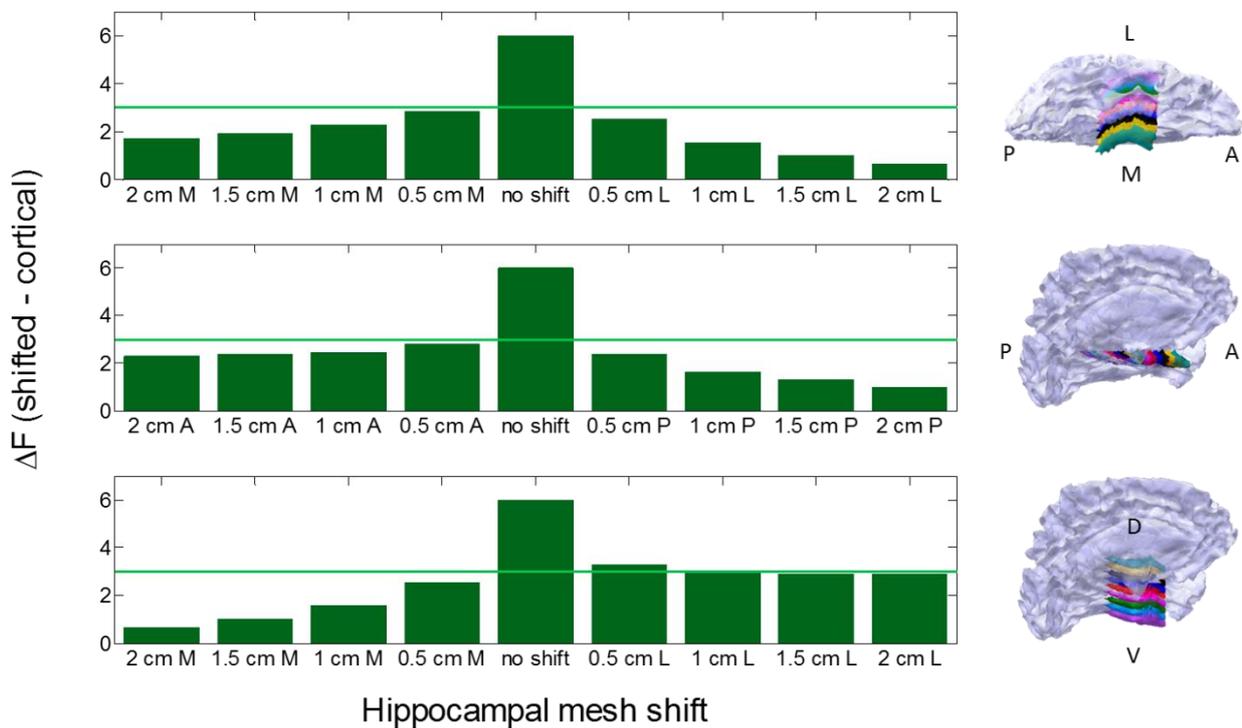


Figure 2.9: Effect of shifting the hippocampal mesh on Free energy.

We compare different combined models with shifted hippocampal meshes to the standard cortical (hippocampus-free) model. Bars represent $\Delta F = \text{mean } F_{\text{shifted}} - F_{\text{cortical}}$ of 30 different hippocampal simulations for the particular shifted model. Top panel shows medial-lateral shifts, middle panel anterior-posterior, bottom panel up-down. While no shift (combined – cortical) gives a significant ΔF value, shifting the hippocampus in any dimension or direction renders the model comparison non-significant. $\Delta F = 3$ is taken as a significance threshold and marked in green.

Dipole Localisation Error

We also performed more traditional analysis by calculating the dipole localisation error (DLE) between simulated and recovered sources (**Figure 2.10**). Both the average DLE and its variance increases as co-registration error increases (A) and SNR decreases (B). Furthermore, we found that in accordance with our Free energy results (**Figures 2.6 and 2.7**), DLE is more affected by co-registration error than by SNR. By definition, DLE can only be calculated when the true source location is known, i.e. in simulations. Critically therefore, the correspondence between the DLE and Free energy supports the notion that Free energy is valid and informative when the true source location is not known, i.e. in empirical data.

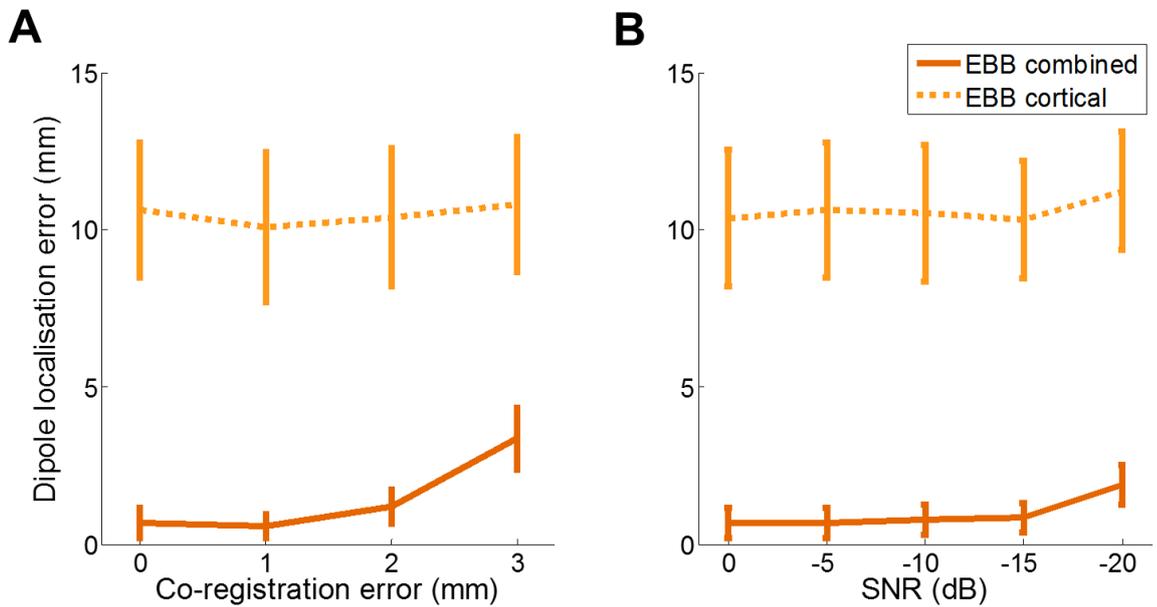


Figure 2.10: Dipole localisation errors as a function of co-registration error and SNR when sources are hippocampal.

A) Mean dipole localisation error (\pm SEM) against co-registration error. SNR of -5 dB. Dotted yellow lines show results for EBB using the cortical model; orange solid lines used for combined. For the combined mesh, DLE and variability starts to increase >1 mm co-registration error. **B)** Mean dipole localisation error (\pm SEM) across SNR levels. For the combined mesh, both error and variability increases with noise >15 dB. No co-registration error added.

Furthermore, we quantified how often the hippocampal simulations have source distribution maxima on the hippocampal mesh (the true positive rate or sensitivity), and how often cortical simulations have maxima on the cortical mesh (the true

negative rate or specificity). At SNR -5 dB and no co-registration error, we find that the sensitivity is 93.33% and specificity is 100%. The full table of sensitivity and specificity values across all co-registration error and SNR levels is shown below:

Table 3: Sensitivity and Specificity values across co-registration error and SNR levels

0 mm co-registration error					
SNR (dB)	0	-5	-10	-15	-20
Sensitivity	93.333	93.333	93.333	93.333	76.667
Specificity	100.00	100.00	100.00	100.00	100.00
1 mm co-registration error					
SNR (dB)	0	-5	-10	-15	-20
Sensitivity	93.333	96.667	93.333	93.333	80.000
Specificity	100.00	100.00	100.00	100.00	100.00
2 mm co-registration error					
SNR (dB)	0	-5	-10	-15	-20
Sensitivity	90.000	86.667	76.667	90.000	76.667
Specificity	100.00	100.00	100.00	100.00	100.00
3 mm co-registration error					
SNR (dB)	0	-5	-10	-15	-20
Sensitivity	80.000	66.667	66.666	56.667	50.000
Specificity	100.00	96.667	100.00	100.00	96.667

Multiple Sources

One further question is whether this approach is robust to situations containing a mixture of cortical and hippocampal sources. Reconstructing concurrent sources in the model comparison framework revealed that even when a single hippocampal source is simulated along with three cortical sources, the model comparison framework can (in some cases) be used to identify the hippocampal activity. **Figure 2.11** shows the relationship between ratio of cortical-to-hippocampal sources and $\Delta F_{\text{anatomical}}$. As expected, the proportion of hippocampal activity correlates with Free energy differences: Four cortical sources (Free energy mean -0.0139, BOR 0.814), three cortical and one hippocampal source (Free energy mean 0.7934, BOR 0.191), two cortical and two hippocampal sources (Free energy mean 1.491, BOR <0.001), one cortical and three hippocampal sources (Free energy mean 1.900, BOR <0.001), four hippocampal sources (Free energy mean 4.366, BOR <0.001). Importantly this analysis also acts as a second validation of the claim that (multiple) purely cortical sources (condition 4C) do not benefit from the addition of the hippocampal mesh (**Figure 2.5B**).

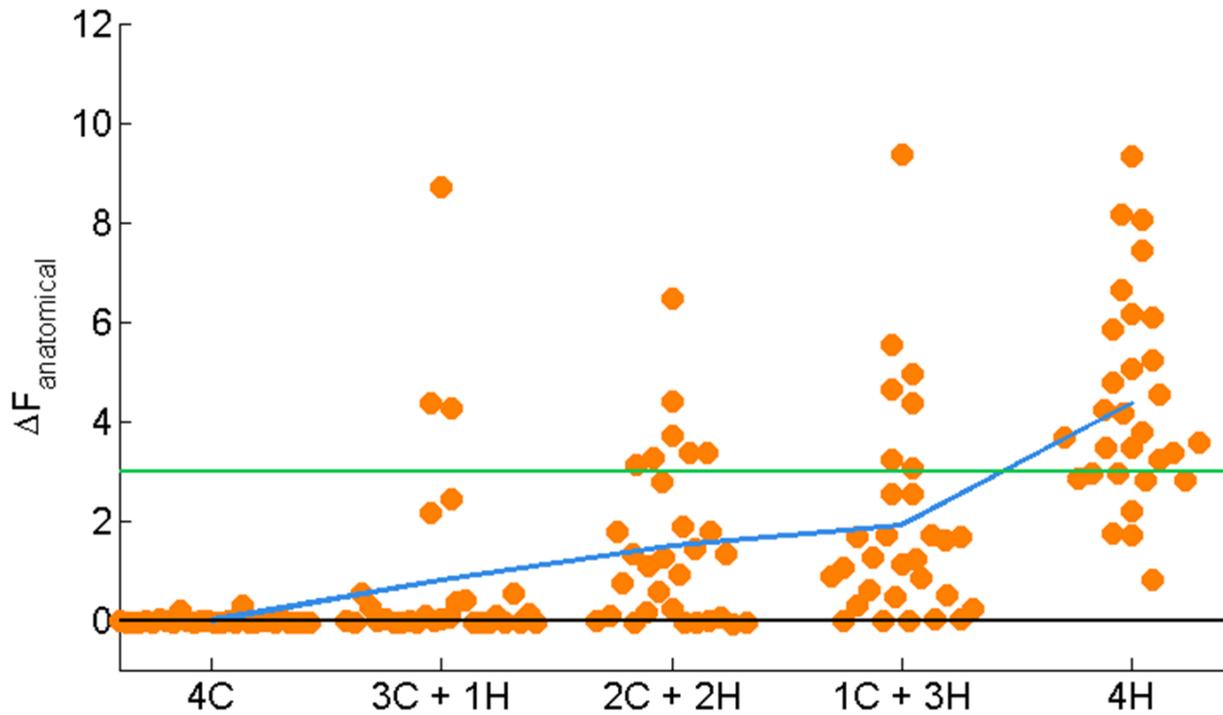


Figure 2.11: Simultaneous sources.

To test whether the model comparison framework would generalise with more than one dipole, we simulated four simultaneous dipoles at different ratios of cortex (C) to hippocampus (H). Orange dots represent 30 $\Delta F_{\text{anatomical}}$ values with 10 root mean square (rms) noise added. Blue line shows mean Free energy difference. We added no co-registration error to these simulations. As the proportion of hippocampal sources increases, the Free energy differences increase. To add noise, we simulated band-limited white noise waveforms between 1-80 Hz for 300 ms. The effective dipole moment for cortical sources was set to 100 nAm and 200 nAm for hippocampal sources (Attal et al., 2012; Murakami and Okada, 2015, 2006). The simulation locations were the same as used previously (which were drawn at random). Each simulated dataset had a sampling rate of 600 Hz with the sensor-level white Gaussian noise level now defined as an absolute value of 10 root mean squared (rms). Due to the range of frequencies simulated, we used 16 temporal modes to model the data. We added no co-registration error to these inversions.

Differences Between Functional Priors

We then asked whether we can use the simulated datasets to directly compare the performance of the functional, as opposed to anatomical, priors. To do this, we looked at the Free energy values obtained with the combined model and compare these

values across algorithms. Given the same data and anatomical models, the differences therefore reflect the appropriateness of the assumptions related to the source covariance, and not, as before the physical locations of potential sources. The $\Delta F_{\text{functional}}$ is defined as the difference between given algorithm and the algorithm which returns the smallest Free energy value (so for the worst algorithm this difference will be zero). **Figure 2.12** shows the mean $\Delta F_{\text{functional}}$ from the 30 hippocampal simulations, across SNR levels (A), and co-registration error levels (B). We found that the most likely functional prior is MSP for all SNR levels and that the second best algorithm is EBB throughout. This implies that EBB has a good empirically-based estimate of source power even at low SNRs and high co-registration error, considering that it does not have the advantage of MSP where a small sub-set of sparse priors pre-specified.

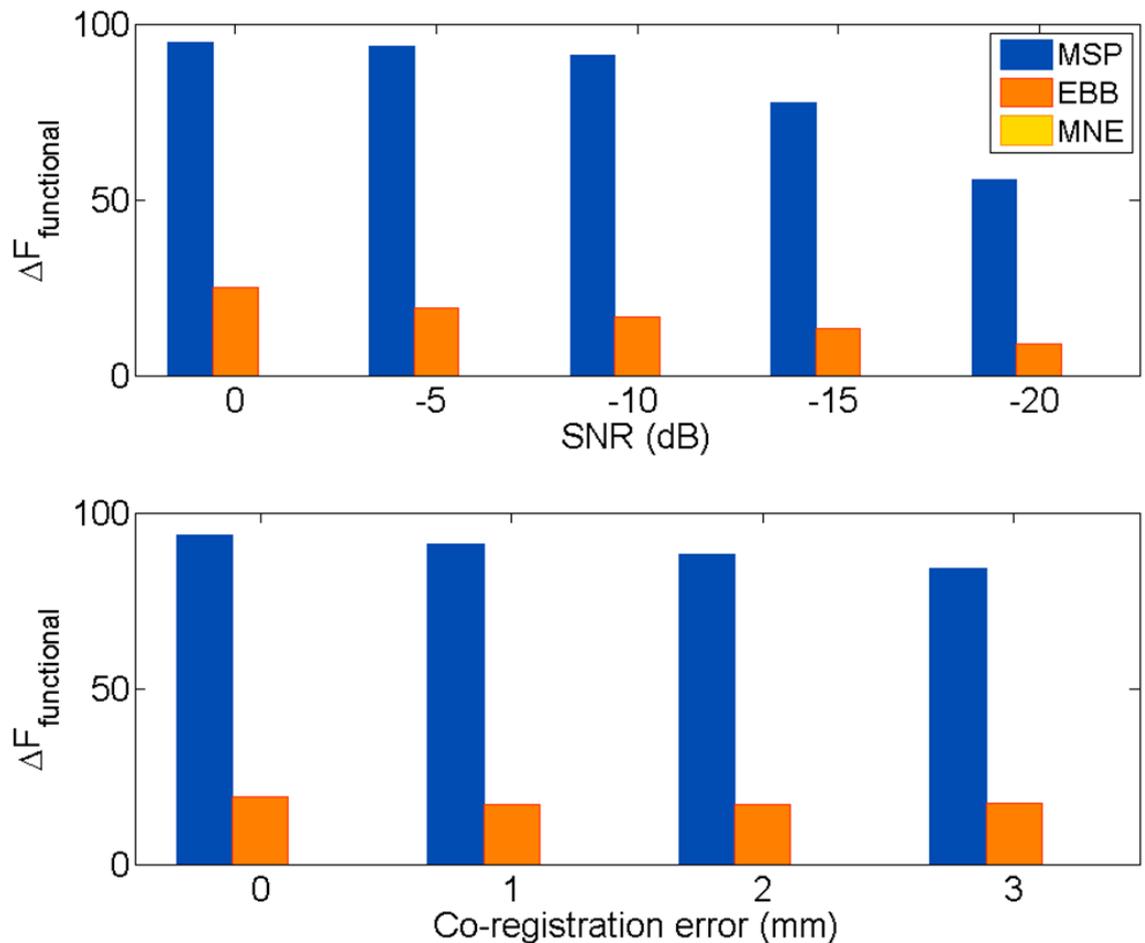


Figure 2.12: Functional model comparison.

Functional model comparison. This analysis compares combined models across algorithms, meaning that we compare the functional, as opposed to anatomical, priors. Given that data and models are constant, the differences in Free energy reflect the appropriateness of the

assumptions related to source covariance. **A)** Model comparisons of combined models across algorithms as a function of SNR. Bars encode mean functional Free energy difference ($F_{\text{algorithm}} - F_{\text{worst algorithm}}$) of 30 hippocampal source inversions. Results are normalized to smallest mean functional Free energy difference at each SNR value. Results are for the combined model without co-registration error. The most likely functional prior is MSP for all SNR levels tested, and the second most likely is EBB throughout. This implies that EBB has a relatively good data-driven estimate of source power, even at low SNR (considering that it does not have the explicit advantage of MSP where a small subset of correct and sparse priors are used to constrain the source space). **B)** Same as A, but as a function of co-registration error. Results are similar to SNR range, implying that the model comparison differences are driven by differences in the model evidence values associated with the combined (compared here), and not the cortical models. SNR set to -5 dB throughout.

Discussion

We demonstrate a new method for making probabilistic statements specifically about hippocampal engagement in MEG. We show that in order to reliably infer hippocampal activity through comparison of two generative models, one with and one without the hippocampus explicitly modelled, uncertainty about the location of the brain relative to the sensors must be less than 3 mm. Notably, this rather stringent criteria applies only when one wishes to make a specific case for hippocampal (rather than medial temporal) involvement.

This approach works because a model without the hippocampus explicitly modelled will be sub-optimal in the sense that it provides a less-parsimonious solution to explain the same amount of data. Consequently, this model will be penalized in terms of its model evidence. Therefore, although the cortical and combined models may explain the same amount of variance in the data (**Figure 2.4B**), the cortical model must use more sources to do so, consequently returning a lower Free energy value (**Figure 2.4C**).

The most immediate advantage of the Bayesian model comparison method is that it allows us to make use of much more information when making the same inference. For example, instead of simply looking at the location of the peak in an image, we can use a generative model to test whether the orientation of the source is what we would have expected.

One important caveat is that our inference is only as good as our models. So for example if the true activity arises from a neighbouring structure (such as the

amygdala) and we have not specifically included a model of the amygdala then we may make incorrect inference. The level of detail required is also an empirical question. It would be interesting to test whether for real data we can distinguish between canonical and individual models of the hippocampus (similar to work on the cortex, see (Henson et al., 2009; Troebinger et al., 2014a)). We are encouraged by the sensitivity of our inference to hippocampal location (**Figure 2.9**). Here we have focused on the distinction between cortical and hippocampal surface sources (**Figure 2.5, Figure 2.8**) but we hope to eventually incorporate structural features of hippocampal subfields and close-by structures (retrosplenial cortex, parahippocampal cortex, entorhinal cortex, amygdala, etc) into the modelling of neuronal current flow. This would allow the uncertainty to be further reduced and for us to begin to distinguish between subcortical structures in MEG and study their real-time interactions.

Although the spatial resolution is inevitably poorer at deep locations in the brain (Hillebrand and Barnes, 2002), we have shown that the approach presented here is sensitive enough to discriminate between hippocampal and neighbouring cortical sources, even when these are as close together as ~2 mm (**Figure 2.8**). We attribute this discriminability to the different orientations of the local surfaces which give us leverage to distinguish between models not commonly available in more traditional voxel-wise inference where only location information can be used. As such, Bayesian model comparison is distinct, and complementary to standard group level voxel wise statistics in which we traditionally look for a peak location within a specific structure. The key difference being that for each subject we have anatomical models which constrain not only source locations but also orientations (and potentially in the future expected current densities (Helbling et al., 2015)) which give us an extra dimension through which to distinguish between models.

With respect to the central question of whether significantly higher Free energy for the combined model is specific to hippocampal activity, we conclude that it is. This is supported by four lines of converging evidence: *a*) Free energy is not higher for the combined model when the source(s) is/are cortical (**Figure 2.5B** and **Figure 2.11**), *b*) simulating activity on the nearest portion of medial temporal lobe does not give rise to significant Free energy differences (**Figure 2.8**), *c*) the maximum Free energy difference is specific to the correct location of the mesh and falls below significance if the mesh is shifted (**Figures 2.9**), and *d*) using the combined mesh, the dipole localisation error is close to zero at low co-registration error and high SNR (**Figure 2.10**). Thus, the extent to which Free energy differences can be used to infer

hippocampal activity is dependent on the accuracy of the solution obtained with the combined model. Largely irrespective of the SNR level, the most important empirical factor when attempting to unambiguously determine the presence or absence of hippocampal activation is minimization of co-registration error. Notably, we base the detection of hippocampal on six cycles of oscillatory activity here (20 Hz simulation frequency and 300 ms duration). Having more data would increase the detectability by improving the SNR (Brookes et al., 2008).

We simulated data using one set of functional priors (suited to MSP) and reconstructed using two other commonly used assumption sets (beamforming and minimum norm). As one might expect, the MSP performs the most robustly and sensitively. This is unsurprising, given that the simulated activity was sparse, a characteristic that matches with the MSP assumptions. Another important point is that we pre-selected the correct set of priors (spatial patches) for MSP and therefore side-stepped a potentially computationally intensive search over possible patches which would be necessary for empirical data (for example see (Troebinger et al., 2014a) where we used 32 random patch sets per dataset and cortical model). This means that while MNE and EBB had the same large search space, only MSP was given priors to start the search from which exactly matched the actual simulation location. Overall we were encouraged to find that all the functional assumptions showed a preference for the correct anatomical model (**Figure 2.3B-D**) and gave somewhat similar estimates of the true source distribution (**Figure 2.4A**). Importantly, as the true functional priors will never be known, the Free energy equation (Equation 1.16) also allows us to select the most likely functional priors (**Figure 2.12**). Given that the EBB algorithm did not have the advantages of the reduced MSP prior space, yet performed well, and given the wealth of previous hippocampal studies using volumetric beamformers (Cornwell et al., 2012; Guitart-Masip et al., 2013; Kaplan et al., 2012b; Poch et al., 2011), we think this is a promising avenue for further work.

It is important to consider the main limitations and assumptions of using Bayesian model comparison and Free energy. Firstly, as is true for any model comparison scheme, we cannot evaluate how good the individual models are in absolute terms; we can only infer how good they are relative to one another. It is therefore not possible to make inferences or predictions about whether alternative models might be better without testing these models. In addition, there is a risk of having local maxima in the cost function (in this case the Free energy) if the number of sources and/or hyper-parameters is very large (Wipf and Nagarajan, 2009). This would mean that models could converge on non-optimal solutions and thereby render the Free energy value an

invalid reflection of the model or algorithm's optimal parameter settings. However, using simulated data, it has been shown elsewhere that Free energy correlates with cross-validation accuracy using machine learning approaches (Penny and Roberts, 1999), and with conventional reconstruction evaluation measures such as dipole localization error (Belardinelli et al., 2012). We also find this in our data (**Figure 2.10**). It follows that maximization of Free energy can be used to fine-tune features of the generative model used for analysis, such as number of equivalent current dipoles (Kiebel et al., 2008), forward model (Henson et al., 2009), or cortical layer giving rise to the measured signal (Troebinger et al., 2014a). However, perhaps the greatest advantage of Free energy is that it provides a framework for reliably evaluating hypotheses without knowledge of ground truth.

Here we have evaluated algorithm performances for a set of specific perturbations from ideal conditions. We emphasize that there are parameters which we have not fully investigated the effects of. For example, it would be interesting to evaluate the algorithms using different types of correlated noise (although see **Figure 2.11** in which correlated noise is effectively introduced through multiple sources). Ultimately, there are therefore still unresolved questions related to the assumptions implicit in the algorithms and simulation parameters used here. Nonetheless, we show that irrespectively of these, source reconstruction of hippocampal activity depends upon accurate co-registration between MRI and MEG data.

The outstanding issue therefore is whether the proposed generative model will be useful in practice. We know from these simulations that the main empirical constraint will be co-registration error which we can now reduce down to <1.5 mm using flexible and subject-specific head-casts for MEG. Moreover, the head-casts reduce head movement during recording to <0.4 mm which gives way to higher SNR data. We are now working on providing empirical validation of the model comparison approach presented using these devices (Troebinger et al., 2014b) conjunction with a paradigm known to modulate hippocampal activity (Doeller et al., 2008).

The roles of the hippocampus in cognition has been emphasized in both humans (for example, Burgess et al., 2002; Lega et al., 2012; Rutishauser et al., 2010; Zhang and Jacobs, 2015) and animals (Kahana et al., 2001; Logothetis et al., 2012). Our work shows that by optimising acquisition protocols such that co-registration error is minimized and SNR is maximised, e.g. by using head-casts (Troebinger et al., 2014b), we have the ability to selectively study hippocampal dynamics in humans non-invasively.

Key points

- We demonstrate a method for quantifying hippocampal engagement probabilistically using simulated hippocampal activity and realistic anatomical and electromagnetic source modelling.
- We constructed two generative models, one which supports neuronal current flow on the cortical surface, and one which supports neuronal current flow on both the cortical and hippocampal surfaces.
- Using Bayesian model comparison, we could then infer for any given dataset which of the two models provided a more likely explanation of the data.
- In addition, we tested the robustness of this inference by adding co-registration and sensor level noise.
- We found that the framework is sensitive to hippocampal activity when co-registration error is <3 mm and the sensor-level signal-to-noise ratio (SNR) is <-20 dB.
- These level of co-registration error and SNR can now be achieved empirically using recently developed subject-specific head-casts.

This chapter derives from Paper 1: “Using MEG generative models to make probabilistic statements about hippocampal engagement” Sofie S Meyer, Holly Rossiter, Matthew Brookes, Mark Woolrich, Sven Bestmann, Gareth R Barnes (under review, NeuroImage).

Chapter 3

Experiment 2: Flexible Headcasts for High Spatial Precision MEG

Precis

In combination with magnetoencephalographic (MEG) data, accurate knowledge of the brain's structure and location provide a principled way of reconstructing neural activity with high temporal resolution. However, measuring the brain's location is compromised by head movement during scanning, and by fiducial-based co-registration with magnetic resonance imaging (MRI) data. Each factor contributes in the order of 0.5 cm of error which is propagated into the forward model. Here, we present a method for stabilizing and repositioning the head during scanning, and co-registering MRI and MEG data with low error. Using this new flexible and comfortable subject-specific head-cast prototype, we find within-session movements of 0.25 mm and between-session repositioning errors around 1 mm. Further, we empirically demonstrate high precision source level reproducibility.

Introduction

In theory, the spatial precision attainable with magnetoencephalography (MEG) increases monotonically with increasing signal strength (Gross et al., 2003; Hillebrand and Barnes, 2005, 2003). In practice however, this increase is difficult to achieve. Two of the main limitations are co-registration between functional MEG data and anatomical magnetic resonance imaging (MRI) data, and head movement during scanning. Both introduce, at best, ~ 0.5 cm of uncertainty about the location of the head relative to the sensors (Adjamian et al., 2004; Gross et al., 2013; Ross et al., 2011; Singh et al., 1997; Stolk et al., 2013; Whalen et al., 2008). Both sources of error non-linearly compromise the forward modelling accuracy (Hillebrand and Barnes, 2011, 2003), and reduce the signal-to-noise ratio (SNR) through topographical blurring (Medvedovsky et al., 2007; Uutela et al., 2001).

Although some progress has been made in minimizing co-registration error (Hironaga et al., 2014; Koessler et al., 2011; Nunez and Silberstein, 2000; Whalen et al., 2008), for example by stabilizing the head during recording (Adjamian et al., 2004; Singh et al., 1997), or compensating for movements both during and after recording (Medvedovsky et al., 2015, 2007; Nenonen et al., 2012; Stolk et al., 2013; Uutela et al., 2001), implementation problems have remained. The sources of residual error include misalignment of surfaces, amplification of small placement errors at the front

of the head to large errors at the back of the head, and/or reliance on invariance in fiducial placement within and across experimenters and subjects (Adjamian et al., 2004).

Using 3D printing to create solid head-casts which are moulded to the surface of the head internally and to the inside of the MEG dewar externally, we recently showed reduction of co-registration errors to <2 mm (Troebinger et al., 2014a, 2014b). Although these first solid head-casts gave access to much higher quality data by minimizing both co-registration error and head movement, they covered the eyes and their rigidity reduced participant comfort, particularly for long recording sessions. Here, we present a new head-cast prototype made of flexible polyurethane foam which leaves the eyes uncovered, and is easier, safer, and more comfortable to use. The improved user comfort is primarily because of the flexibility which makes it easier and faster to get into and out of the MEG helmet. Furthermore, the 3D printing is now based on an MRI image (as opposed to an optical scan used in Troebinger et al., 2014a and 2014b) which both maximise the accuracy with which the cast fits the head, and minimizes co-registration error by predefining the MEG fiducial coil locations in MRI space. We describe the construction pipeline, the within- and between-session head movement for subjects wearing these head casts, and assess the estimated co-registration error. We then show how these improvements give rise to very high between-session reproducibility at source level.

Hypothesis and objectives

Hypothesis: If co-registration error and movement of the head during data recording can be minimized, then spatial precision and data reproducibility will be maximised. If subjects can be re-positioned consistently, and maintain a stable head position throughout normal-length scanning sessions, then it will be possible to build up high-SNR datasets through repeated scanning of single subjects. If this is true, then non-varying electrophysiological responses should be consistent across re-positionings and scanning days.

In addition to maximising data quality through stabilization and reliable re-positioning of the head relative to the sensors, the objective of this chapter is to design the head-casts such that subject comfort and safety are improved relative to the first head-cast prototype.

Methods

This section is divided into two parts. First, we describe the methods used for building head-casts. Next, we describe the scanning procedures for evaluating the head-casts with respect to head stabilization, co-registration, and spatial precision.

Participants

Data were collected from five healthy adult subjects (5 men, mean age 30.0 years old). All subjects were right-handed and had no history of neurological or psychiatric disease. One participant was excluded from the analysis because of recording errors. Informed written consent was given by all subjects prior to scanning and the experiments were carried out after obtaining ethical approval from the University College London ethics committee (ref. number 5833/001).

MRI Data Acquisition

In order to construct the head-cast, an accurate image of the scalp surface is required. To get this, we first scanned participants in a magnetic resonance imaging (MRI) system (**Figure 3.1a**). Images were acquired using a Siemens Tim Trio 3T system (Erlangen, Germany). During the scan, the participant lay in the supine position with their head inside a 12-channel coil. Acquisition time was 3 min 42 s, plus a 45 s localizer sequence. We were very cautious of skin distortions as any such errors could potentially make the head-cast ill-fitting and therefore uncomfortable. For this reason, participants were not given padding or ear phones, as these could displace the skin on the face, head or neck. To minimize audible noise they were instead given ear plugs. The short acquisition time minimizes motion and potential consequential distortions. We used an radiofrequency (RF) and gradient spoiled T_1 weighted 3D fast low angle shot (FLASH) sequence with the following acquisition parameters: image resolution 1 mm^3 (1 mm slice thickness), field-of view set to 256, 256, and 192 mm along the phase (A–P), read (H–F), and partition (R–L; second 3D phase encoding direction) directions respectively. Susceptibility differences existing at air-tissue interfaces can lead to magnetic field inhomogeneity and subsequent distortions or signal loss in the acquired image. Therefore, to preserve brain morphology we used a

single shot approach with high readout bandwidth (425 Hz/pixel) and minimum echo time (2.25 ms). Consequently no significant geometric distortions were expected or observed in the images. A short repetition time (7.96 ms) was used to minimise acquisition time while the excitation flip angle was set to 12° to ensure sufficient signal-to-noise ratio for the resulting anatomical image. To accelerate the acquisition, a partial Fourier (factor 6/8) acquisition was used in each phase-encoded direction.

Head-cast Construction

The construction process can be divided into seven steps (**Figure 3.1a-g**). First, we extracted the scalp surfaces from the MRI data using standard SPM12 procedures (<http://www.fil.ion.ucl.ac.uk/spm/>) (**Figure 3.1a**). We then converted this tessellated surface into standard template library (STL) format (**Figure 3.1b**), commonly used for 3D printing. To specify the shape of the fiducial coils, we used optical white light scanning to obtain a 3D representation of a single coil. This was digitally drawn in 3D and then checked for its accuracy both against the digital white light scan as well as the physical coil, using digital measuring callipers. Next three copies of this virtual coil were placed, as per convention, at the approximate nasion, left peri-auricular (LPA), and right peri-auricular (RPA) sites. Note that this was not strictly necessary as any set of distant scalp locations would have enabled the co-registration procedure. This approach therefore does not suffer from inaccuracies in determining anatomical landmarks, as is commonly the case when placing fiducial coils on the head during MEG data acquisition. One constraint on the placement of the coils was ensuring that the coil-body and extruding wire were flat against the scalp, in order to remove unnecessary stress or movement of the coil when the head-cast was put on or taken off.

The original design (Troebinger et al., 2014b) was altered so as to now include eye-hole extensions, ear flaps which extend down below the ears, and a top spacing-cylinder to accurately position the positive head model in the dewar-helmet (**Figure 3.1c-f**). The ear flaps facilitate getting into and out of the scanner more easily and safely (see *Safety Procedures* for more details) and also provide an external reference of when the head-cast is touching the top of the dewar. The virtual 3D model was thus

placed inside a virtual version of the scanner dewar-helmet (**Figure 3.1d**) such that the distance to the sensors was minimized (by placing the head as far up inside the dewar as possible) while ensuring that vision was not obstructed. Next, the positive head-model (plus spacing elements and coil protrusions) was printed using a Zcorp 3D printer with 600 x 540 dots per inch resolution (**Figure 3.1e**). The 3D printed head model was then placed inside the manufacturer-provided replica of the dewar-helmet and liquid resin was poured in between the surfaces to fill the negative space. The resin expands and sets within ~30 s, and the resulting flexible foam constitutes the subject-specific head-cast (**Figure 3.1f**). Note that the coil protrusions on the 3D print now become indentations in the foam head-cast. The fiducial coils can thus be placed inside the resulting indentations and the head-cast can be worn for scanning (**Figure 3.1g**). This removes inaccuracies in determining anatomical landmarks for fiducial placement, and also ensures that the same location is used for repeated scans.

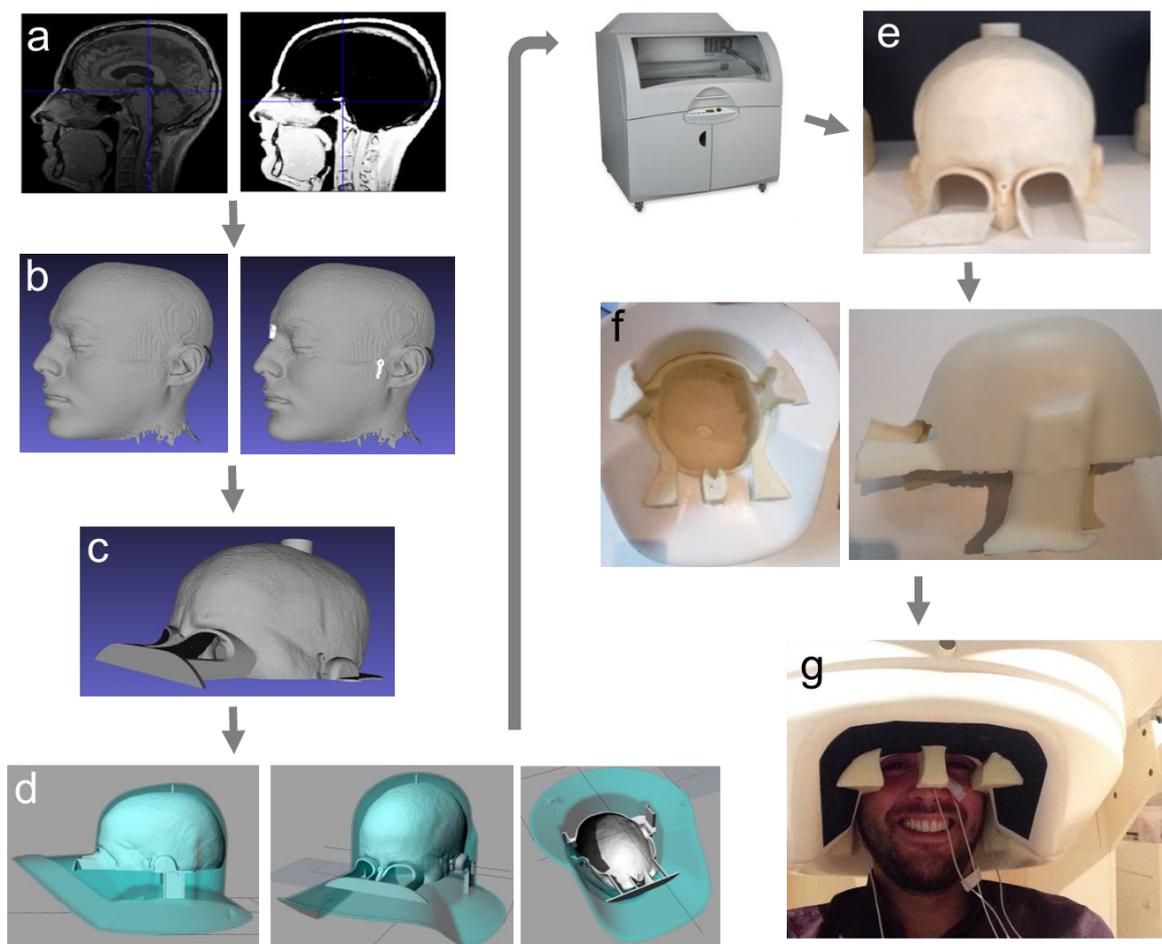


Figure 3.1: Overview of head-cast construction steps.

a) Head surface is extracted from an anatomical MRI image using the standard SPM12 segmentation procedure. **b)** Head surface extraction is converted to a surface file and fiducial coils are added. The coil locations are defined in MRI coordinates. **c)** A positioning cylinder at the top of the head is added to the virtual model to define the position of the head inside the head-cast. Eye extensions are added to enable vision during use. **d)** Using and adjusting the positioning cylinder, eye extensions and ear extensions, the virtual head model is positioned appropriately inside a virtual copy of the MEG dewar. **e)** The positive head model is 3D printed. **f)** The 3D print is placed inside the manufacturer-provided dewar copy (as in d) and foam resin is poured in to fill the gap between the printed positive head model and the dewar. The fiducial coil protrusions on the 3D printed head result in small coil-shaped indentations in the head-cast (the nasion coil is visible between the eye protrusions in the image). **g)** The subject can now wear the flexible foam head-cast and enter into the (real) MEG dewar for scanning.

MEG Data Acquisition

MEG recordings were made using a 275-channel Canadian Thin Films (CTF) MEG system with superconducting quantum interference device (SQUID)-based axial gradiometers (VSM MedTech, Vancouver, Canada) in a magnetically shielded room. The data collected were digitized continuously at a sampling rate of 600 Hz. We refer to *Safety Procedures* for a description of the general operating and safety procedures.

Experiment 1: Between-session variability

We first tested how consistently subjects could be repositioned within the MEG scanner by asking them to reposition themselves in the scanner ten times. In addition to measuring absolute location of the head-cast using the fiducial coils, we also placed a reference coil on one side of the nose to measure relative displacements between the head-cast and head. Each subject performed ten separate 10s trials. For each run, the subject first positioned themselves inside the scanner with the head-cast on, sat still for 10s, before and after which the fiducial coils were localized, and the subject then exited the scanner and removed the head-cast. This removal and replacement was repeated ten times.

In addition to the healthy subjects, we also performed a similar experiment using the manufacturer provided spherical current dipole phantom. This experiment was done in order to get an approximation to the system-based noise inherent in localization of the fiducial coils and for comparison with the head-cast results. We did not have a

head-cast for the phantom but kept the four fiducials fixed on the surface of it using tape. To mimic the re-positioning, we physically shifted its location between ten 10 second trials.

Experiment 2: Within-session variability and button presses

To test the head movement within trials, we analysed head-position data from a single subject (subject 3 from experiment 1) performing button presses across twelve 15-minute sessions with 180 trials each. These sessions were spread over four days (which were separated by several weeks) with three runs per day. Visual stimuli consisted of dots moving left or right, with the participant responding with a button press using the right hand, upon a subsequent Go signal. MEG data were acquired at a sampling rate of 1200 Hz.

Data were epoched around the button press onset (time 0), and a beamformer covariance matrix constructed based on the data from the beta band (15-30 Hz) from -2000 to 2000 ms. To extract the source locations, beamformer-based volume-of-interest (VOI) analysis was then carried out, comparing two time windows ([-1500 to -1000] versus [500 to 1000] ms) to generate a statistical chi square volume centred on the average left primary motor cortex peak (-34, -30, 52 mm in MNI space) with a 20 mm radius and 1x1x1 mm³ grid resolution. The data were subsequently smoothed with a full-width half-maximum kernel of 8 mm. We then constructed a time frequency decomposition of the signal from the primary motor cortex sphere (centred around -34, -30, 52 with 20 mm radius) using a Morlet wavelet transform method with 7 cycles on the baseline ([-1500 to -1000] ms) corrected data.

Results

Between-session movement

To first establish how reproducible the absolute head position was when using head-casts, we measured the fiducial coil locations across ten repositioning trials (Experiment 1). We found that it was possible to reposition the fiducial coils relative to the MEG system within 0.6 mm standard deviation in any one dimension (**Figure 3.2a**).

Next we were interested in whether there is a risk of the coils moving with respect to each other when the head-cast is taken on and off. We examined this by calculating the standard deviation of the distances between fiducial coils across repositioning trials. We found no such effect measurable as the standard deviations of the distances were similar to the standard deviation of the absolute locations (**Figure 3.2b**). We found that when we repeated the experiment using a phantom (with the coils fixed on the surface), we observed a similar level of variability, suggesting that this error is due to uncertainty in the (MEG system's) localization of the coils themselves and not to coil movement.

Since the fiducial coil locations are recorded by the MEG system, changes in head-position relative to the dewar during recording, although undesirable, can be accounted for. A more pernicious source of error is relative movement of the head with respect to the head-cast. To address this directly, we placed a reference coil on the nose of the subject in order to measure the distances between this reference and the standard fiducial coils (**Figure 3.2c**). Unlike with the previous analysis where there was no difference between measurements made with the phantom and normal subjects, we now observed an effect beyond measurement error. We found that the variability in the location of the head-cast relative to the head was predominantly due to uncertainty in the Z dimension of 1.2 mm standard deviation.

Next, we were interested in whether these differences in distances to the reference coil could be attributed to differences in location along some spatial dimensions more than others. **Figure 3.2d** shows that the most variable dimension is the Z (up-down) dimension. **Figure 3.2e** shows the standard deviation of the reference coil with respect to 'head-centred' space, meaning that the coordinate frame is defined by the three standard fiducial coils. These values reflect how much the reference coil moved around relative to the standard fiducial coils inside the head-cast in X (front-back), Y (left-right), and Z (up-down) dimensions. We thus found that the main axis along which additional variance occurs is the Z (up-down) axis (**Figure 3.2d,e**). Surprisingly, we found this highest variation in the Z dimension to be true for both phantom and human measurements. This suggests increased measurement uncertainty in this plane, which may be unrelated to the head-cast but perhaps due to the MEG sensors and algorithms used to localise the coils or simply the vertical movement of the scanning chair (on which the phantom rested) over time.

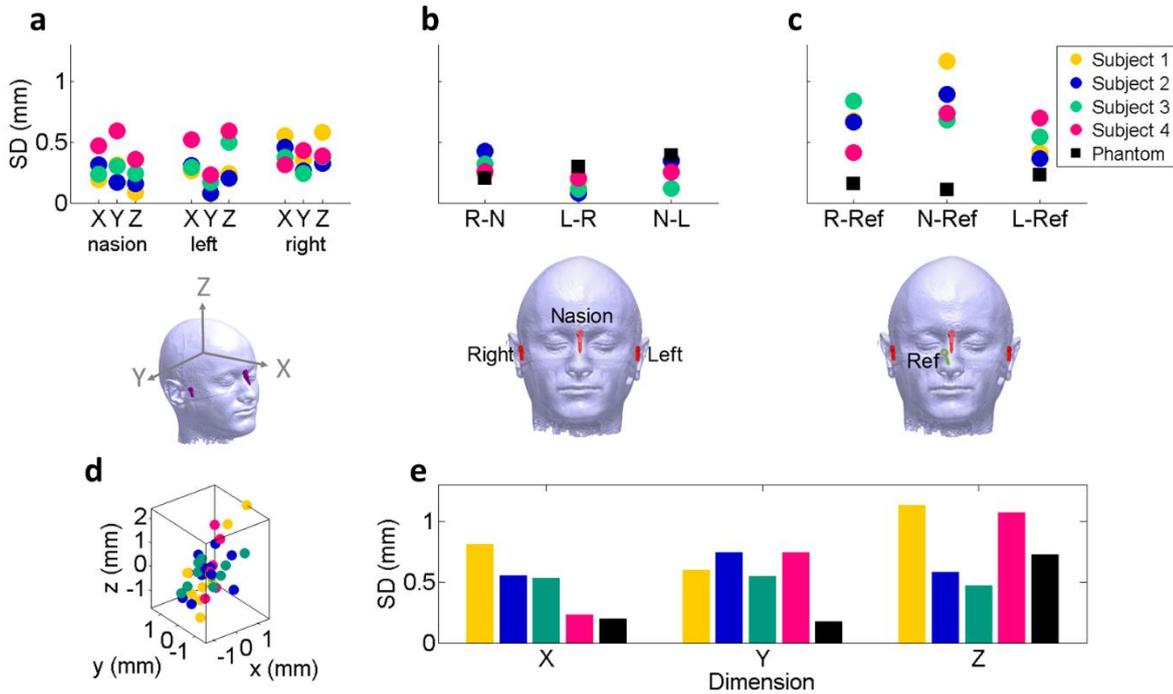


Figure 3.2 Between-session head movement.

Results from Experiment 1 (re-positioning trials where each of the four subjects came out of the scanner, removed the head-cast, put it back on and re-entered 10 times). **a)** Variability of absolute coil locations. Dots show the standard deviation of the absolute coil location over the course of the experiment. Repositioning is precise to within <0.6 mm standard deviation for any coil in any dimension. **b)** Coil-coil distance variability. The standard deviations are calculated from the distances between the fiducial coils measured in Experiment 1. The distances vary <0.5 mm which is within the range of measurement error, as illustrated by the phantom measurements (black squares). **c)** Reference coil-standard coil distance variability. Same format as b, but based on the distances between each of the three standard fiducial coils and a reference coil placed on the nose. There is more variability with normal subjects than phantom. **d)** Scatter plot showing absolute locations of reference coil in head-centred (standard coil-defined) space. This plot illustrates dimensions along which the reference coil location varies relative to the standard coils: mostly in the Z (up-down). **e)** Location of reference coil in head-centred space. Bars encode standard deviation of absolute position of the reference coil in head-centred space measured across 10 repositioning trials. The location of the reference coil deviates <1.2 mm from the fiducial coils in the worst case. Note that variability along the Z dimension is also relatively high with the phantom. The standard deviation over all subjects was 0.50, 0.57, and 0.80 mm for the X, Y and Z dimensions respectively.

In addition to the standard deviation values shown in the Figure, Table 4 shows the mean, maximum and minimum values of the coil locations. Values are in cm. P phantom, S1 subject 1, LPA left pre-auricular, RPA right pre-auricular.

Table 4: Mean, maximum and minimum values for coil locations during Experiment 1

		Nasion			LPA			RPA			Reference		
		x	y	z	x	y	z	x	y	z	x	y	z
S1	Mean	7.17	6.88	-22.56	-4.570	7.11	-24.510	7.03	-4.44	-24.64	6.39	7.97	-26.6
	Max	7.21	6.94	-22.48	-4.53	7.13	-24.43	7.04	-4.39	-24.57	6.44	8.06	-26.4
	Min	7.14	6.85	-22.62	-4.62	7.1	-24.55	7.01	-4.47	-24.76	6.33	7.78	-26.76
S2	Mean	7.86	7.57	-24.76	-3.79	6.75	-24.86	6.25	-3.74	-24.64	6.24	7.71	-26.41
	Max	7.90	7.62	-24.72	-3.77	6.76	-24.83	6.28	-3.72	-24.59	6.30	7.82	-26.29
	Min	7.81	7.54	-24.84	-3.81	6.74	-24.91	6.23	-3.77	-24.69	6.15	7.64	-26.54
S3	Mean	6.88	6.52	-23.46	-4.28	6.64	-24.68	6.55	-4.18	-24.57	6.07	7.38	-26.92
	Max	6.91	6.57	-23.4	-4.230	6.67	-24.64	6.60	-4.07	-24.53	6.16	7.45	-26.85
	Min	6.83	6.47	-23.53	-4.33	6.62	-24.72	6.52	-4.24	-24.64	6.01	7.30	-27.01
S4	Mean	7.43	7.33	-24.03	-3.62	6.58	-25.37	5.85	-3.98	-24.67	5.86	7.71	-27.48
	Max	7.52	7.43	-23.97	-3.51	6.63	-25.28	5.90	-3.88	-24.61	5.96	7.78	-27.36
	Min	7.36	7.23	-24.07	-3.69	6.55	-25.42	5.79	-4.08	-24.73	5.70	7.67	-27.59
P	Mean	5.24	5.50	-27.87	-5.180	5.48	-26.33	5.44	-4.97	-26.72	-4.92	-4.95	-25.03
	Max	7.34	7.86	-27.74	-2.83	6.90	-25.63	7.13	-1.41	-26.22	-1.97	-3.17	-24.89
	Min	1.55	3.23	-28.02	-7.55	2.5	-26.84	3.84	-6.90	-27.5	-6.24	-6.81	-25.29

Within-session movement

To evaluate the head location stability over time, a single subject was scanned on 12 separate trials lasting 15 minutes each (Experiment 2). We found that results were almost identical across fiducial coils. For any coil, relative movements over twelve 15-minute runs were sub-millimetre (<0.75 mm) and the movement predominantly occurred as drift in the vertical direction (left coil shown as an example, **Figure 3.3a**). Note that these traces were mean-corrected (such that the average head position over each 15 minute period was set to zero) but that the standard deviations of these means were 0.25, 0.25 and 0.26 mm for the X, Y and Z dimensions respectively. Across all coils, we found the standard deviations of locations over time to be below 0.22 mm for

any coils in any dimensions (**Figure 3.3b**). The maximal absolute changes in the coil locations were 0.69, 0.5 and 0.75 mm for the left, nasion, and right fiducial respectively (the corresponding minimal changes were 0.06, 0.11, and 0.06 mm). All of the maxima were in the Z (up-down) dimension. We reason that the explanation for the slightly larger absolute changes and standard deviations in this dimension is that the height of the head-cast inside the dewar may change slightly over the course of a trial, e.g. because the subject relaxes and therefore slouches and loses posture more. We also suspect that there is slightly lower sensitivity in the Z axis (see phantom data in **Figure 3.2e**) which could be due to the sensor configuration (see Discussion).

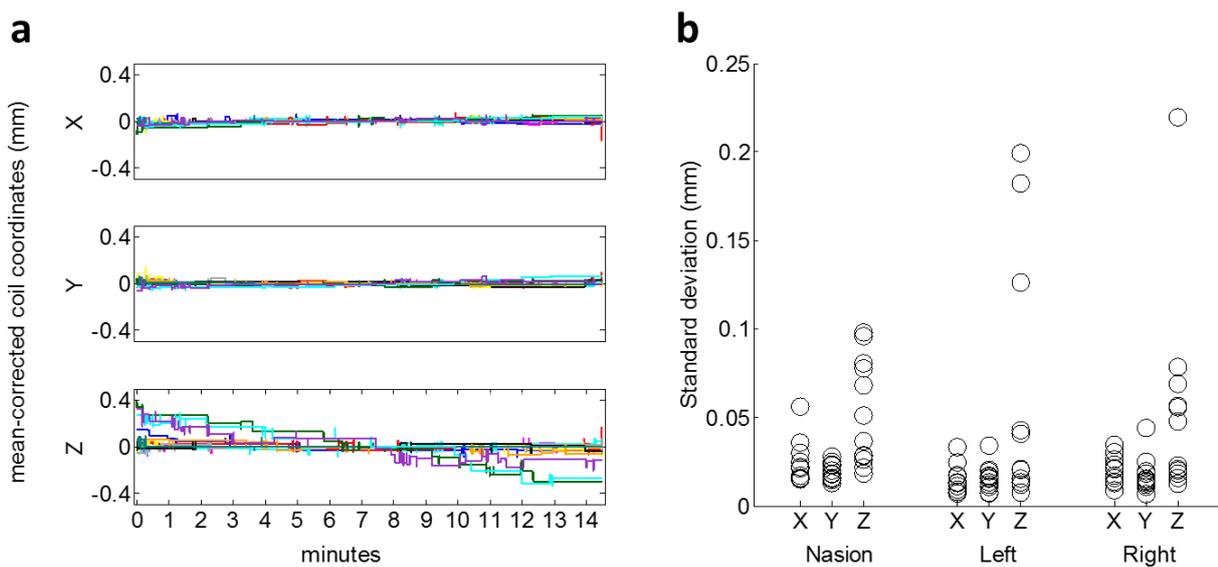


Figure 3.3 Within-session head movement.

Data from Experiment 2. **a**) Absolute location of the left coil in the X, Y and Z dimensions over the course of 12 (colour coded) 15-minute trials. The location is mean-corrected individually for each trial. We find that the variability across time is negligible. The largest movements are downwards (from positive to negative) in line with the subject sliding down in the chair. **b**) Circles show the standard deviations of the absolute coil locations for all 12 trials in all dimensions and for all coils. The standard deviation of the locations recorded was 0.22 mm at maximum. Z (vertical) is consistently the most variable dimension.

Data Reproducibility

In **Figure 3.4** we show recordings from a single subject performing repeated right hand button presses over multiple sessions conducted over several days. The beamformer

peak from 11/12 sessions (consisting of 180 trials each) fell on the same three 1mm³ grid locations (positional noise is added to the plot in **Figure 3.4a** for visualisation) while one fell more dorso-laterally when constrained to the same contralateral hemisphere as the others.

Figure 3.4b shows the time-frequency plot over 4 separate scanning days for the same subject (each represents the average of three 15-minute runs with 180 trials per run) extracted from the beamformer peak in the motor cortex (**Figure 3.4a**).

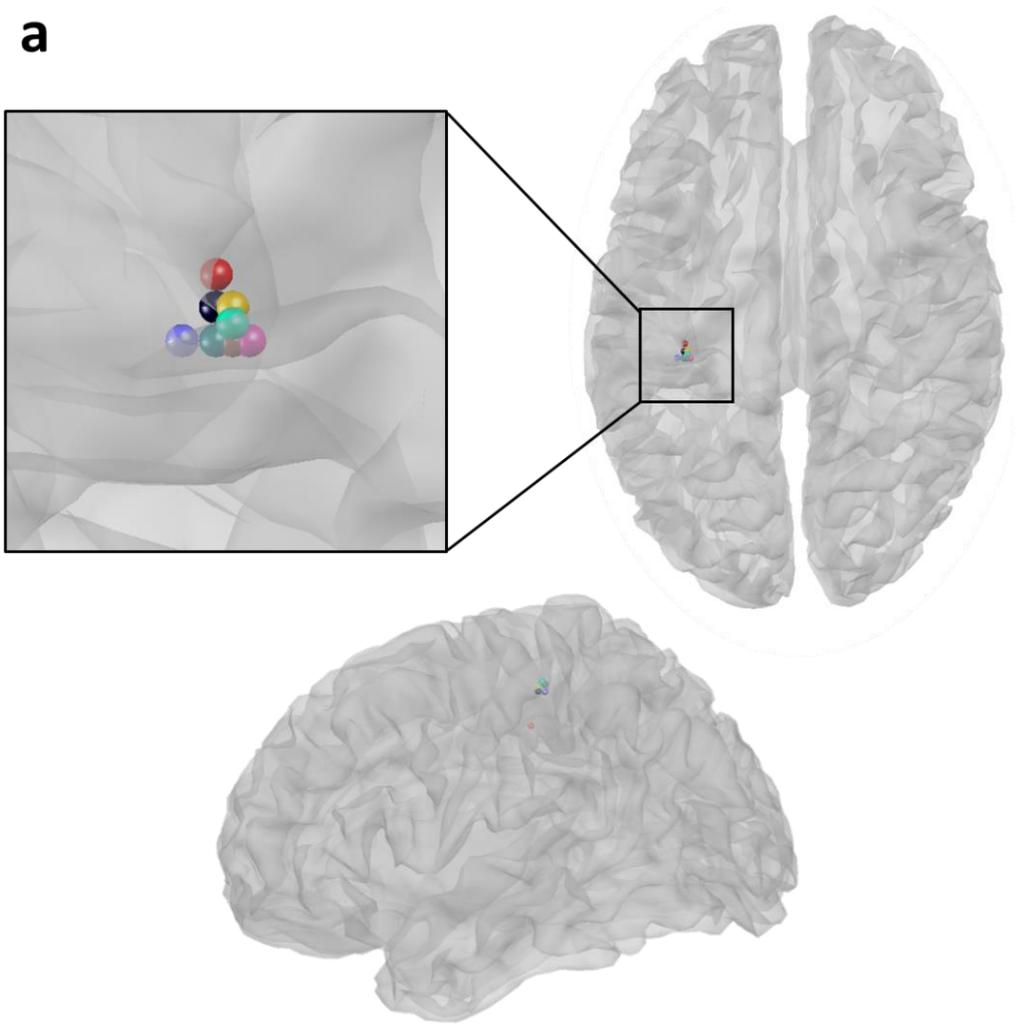
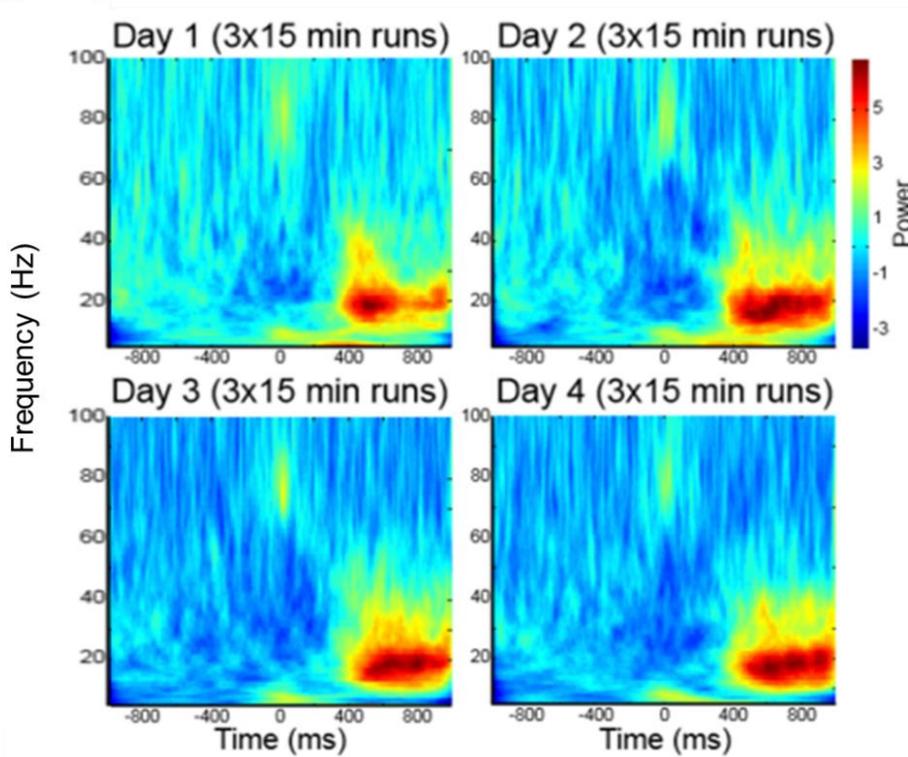
a**b**

Figure 3.4 Consistency of data features across four separate scanning days

a Coloured spheres represent beta (15-30 Hz) rebound peaks from Experiment 2. The peak locations reflect the maximum chi square statistic when comparing pre-button press data (-1500 to -1000 ms) to post-button press data (500 to 1000 ms) across a 20 mm radius, 1x1x1 mm³ resolution sphere centred around the average left primary motor cortex peak (-34, -30, 52). Note that the solutions were not constrained by the mesh as reconstruction was volumetric. Data shown are smoothed using an 8 mm kernel. **b** Time-frequency plots based on the motor cortex beamformer peak for right hand button presses. Each plot shows data from one of four scanning days and represents the average of three separate 15 minute runs, each run consisting of 180 trials.

Discussion

We have developed a novel method for building flexible and subject-specific MEG head-casts to stabilise the head during recording. This method makes use of the subject's MRI image both to build the head-cast by 3D printing an image of the head shape, and to co-register the MEG and MRI data. We find that using this technique for head-cast design, the within-session head movement is 0.75 mm in the worst case (and 0.06 in the best) over a 15 minute period, and the co-registration error is around 1.2 mm.

The head-casts were designed to improve both subject comfort and safety. By making the casts flexible and adding ear flaps, we made it easier to enter and exit the dewar, minimizing the risk of getting stuck or requiring assistance. Additionally, we added eye holes which enable subjects to see and therefore participate in experiments using visual stimuli and/or eye tracking. Together, these features make the head-casts less intimidating to wear and open up the possibility of a wider range of experiments. Importantly the head-cast does not obstruct breathing, vision, or talking although hearing may be mildly compromised. We have not found these head-casts to induce anxiety or claustrophobia.

The other major difference between this generation of head-casts and the previous, is that the 3D print is now based directly on the MRI image eliminating the need for optical scanning. We optimised an acquisition sequence to eliminate distortions on the surface of the head. The manufacturing process is nonetheless not completely straightforward. Whilst some head-casts fit very well, others require removal of

sections that constitute pressure points on the head, typically near the eye holes and temples. This seems to be more pronounced in subjects with longer hair.

With respect to the subjective experience of wearing the head-cast, we find that subjects experience them as constraining and unusual at first, but that they quickly get accustomed to the experience (after a few recordings), and they improve at entering and exiting the dewar. Multiple subjects have remarked that it is obvious to them when the head-cast is fit incorrectly when entering the dewar but not necessarily before. We have also observed that some experienced subjects find it easier to relax while being scanned when wearing a head-cast as they do not have to minimize or inhibit movement. This is an important improvement, as previous methods have relied on self-stabilization (e.g. with bite bars to hold the head in position (Adjamian et al., 2004; Muthukumaraswamy, 2013; Singh et al., 1997)) which induces a risk of increased muscle activity and concomitant artefacts (Kumar et al., 2003; Muthukumaraswamy, 2013; O'Donnell et al., 1974; Whitham et al., 2007).

The main advance of this head-cast approach is that unlike other co-registration minimization approaches, the specification of fiducial points, and extraction of scalp surface based on the same original MRI scan simultaneously minimizes co-registration error and head movement. In turn, this improves the reproducibility of data (**Figure 3.4a**). In previous work (Troebinger et al., 2014b) we have shown that the reduction of within-session movement from 5 to 1 mm gives rise to an effective 5 fold increase in SNR. Notably, high reproducibility implies high precision but not necessarily accuracy. However, the high SNR recordings mean that this framework can be used to directly test between different forward models (e.g. the head in different positions, see Lopez et al. 2012) delivering an accuracy measure that encompasses the complete source reconstruction pathway.

A number of caveats remain. First, we address the increased uncertainty of coil localisation in the Z dimension as observed with increased error in phantom measurements (**Figure 3.2e**). This could either be due to the internal algorithm used to locate the fiducial coils based on their magnetic signature or simply the movement of the scanner-chair. Second, the co-registration estimate based on the reference coil (**Figure 3.2c**) may have been pessimistic as the tape holding the reference coil in place on the side of the nose extended beyond the coil and was easily tugged on by the head-cast. Additionally, the location of the reference coil was both below and outside of the dewar, meaning that it would provide a further challenge to the internal MEG coil localization procedure. Moreover, prospective motion correction methods

where a small optical marker is tracked with sub-micron movement and sub-degree rotation precision has shown that placing the marker on the bridge of the nose is unstable, as uncorrelated movement between the marker and the brain can be observed, likely due to malleability of the skin (Todd et al., 2015).

As mentioned perhaps the most pernicious source of error due to these devices is movement of the subject's head relative to the head-cast. In this case the fiducial locations would appear stable over time whilst, for example, the subject was slowly slipping out of the cast. Based on our reproducibility measurements in **Figure 3.2c** the refitting of the cast over time does not seem to be a problem, but there may be some subjects (due to the shape of their heads) who can slide downwards within the headcast without head-cast movement. In the future we will begin using a 4th coil (attached to the head) for more routine measurements in order to quantify this.

Given that the brain is suspended in corticospinal fluid inside the skull, it must be acknowledged that it remains ambiguous whether the difference between the brain location while supine (during the MRI scan) and sitting (during the MEG scan) could be affecting our estimates. There is a risk that when the head changes orientation with respect to gravity, the brain shifts when the density or thickness of the CSF layer between the brain and the skull changes. It has been approximated that this change in thickness is ~30% which equates to approximately 1 mm (Hill et al., 1998; Rice et al., 2013). We emphasize however that using head-casts while subjects are supine removes the ability to use gravity to exit the dewar, causing the safety to be compromised. Although it would be interesting to directly quantify these shifts through such comparisons, we decided not to due to the safety issues outline below.

Other potential data acquisition problems which we posit that the head-casts solve to a degree but which we have not formally tested are to muscle artefacts (Muthukumaraswamy, 2013), particularly when using bite-bars (Adjajian et al., 2004), and slow within-session drifts (Stolk et al., 2013).

Moreover, we have extended the prototype design such that it can accommodate subject with long or thick hair (Supplementary **Figure 3A**). This extends the usefulness of these devices and means that a larger segment of the population can be scanned. We are working on testing whether this modification affects head stabilization, re-positioning, or in any way introduces unknown errors.

The results of the present study suggest that employment of the individual flexible head-casts for MEG recordings provide an accurate and reliable method of safely

stabilizing the head location during MEG recordings, and for co-registering MRI anatomical images to MEG functional data. This design is ideally suited for studies which require sensitive longitudinal MEG measurements.

Key points

- A method for constructing flexible head-casts to stabilize the head during MEG scanning is proposed
- Co-registration error is minimized by using MRI images to pre-define fiducial coil locations.
- Within- and between-session movement is $<.25$ and <1 mm respectively.
- This enables high reproducibility of source level results.

This chapter derives in part from: Paper 2: “Flexible headcasts for high spatial precision MEG” Sofie S Meyer, James Bonaiuto, Mark Lim, Luzia Troebinger, Holly Rossiter, Sheena Waters, David Bradbury, Simon Little, Sven Bestmann, Matthew Brookes, Gareth R Barnes (submitted, Journal of Neuroscience Methods).

Safety procedures

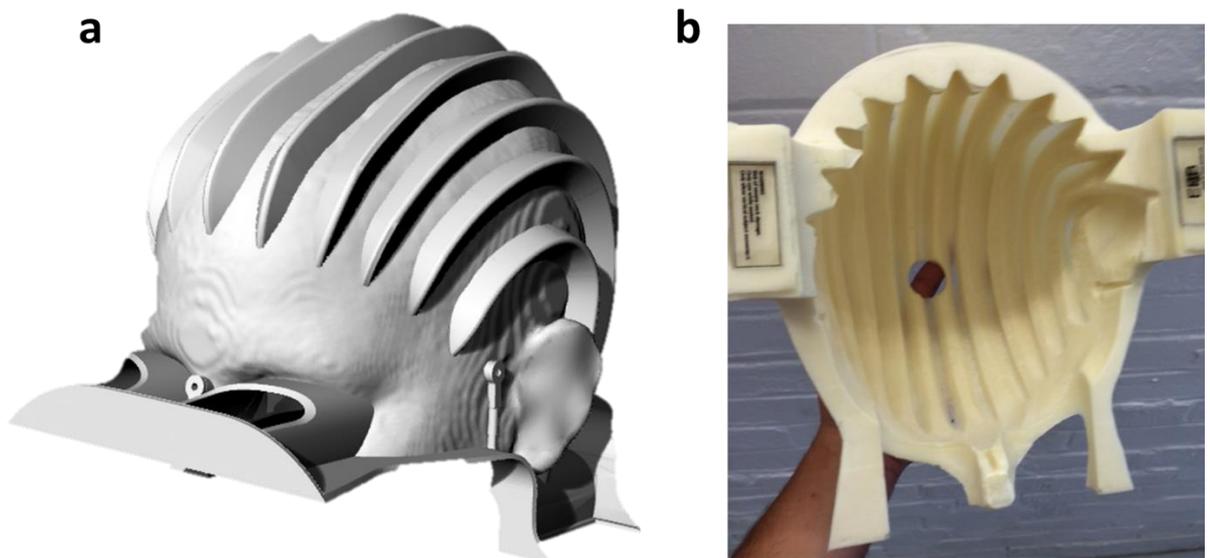
Any head-casts pose a significant source of risk of injury to subjects if used incorrectly. Because the head-casts are designed to fit the subject’s head internally and the MEG dewar externally, the participant’s head is firmly fixed inside the dewar during scanning. This means that any unexpected movement of the chair or MEG system has the potential to cause severe neck injury. Our primary safety measure is therefore to ensure that neither the chair nor the dewar is moved while the subject is wearing a head-cast. This means that the initial positioning of the subject (as well as any subsequent adjustments to the height or angle of the chair) only takes place when the subject is not inside the scanner or wearing a head-cast. To enter or exit the dewar, the subject therefore slides in and out of the seat unassisted. In our experience, this takes some practice but is easily and quickly mastered. However, this means that only healthy, agile volunteer subjects are suitable for head-cast scanning. In order to ensure maximal comfort and safety of participants, we have developed a set of safety procedures to be followed by all researchers carrying out MEG scans involving head-

casts. We also screen subjects to avoid scanning participants with claustrophobia, and place a panic button inside the magnetically shielded room should the subject wish, at any time, to stop scanning.

We advise that only authorised personnel are allowed to scan volunteers with a head-cast.

For these reasons we have decided never to use the head-cast with a subject in supine position where the consequences of unexpected relative movement between the dewar and the bed could be much more serious.

We refer to our safety guidelines, standard operating procedures, training guide, volunteer guide, and emergency procedures available on the MEG community website (<http://megcommunity.org/> under instrumentation > peripherals > subject stabilization) which also contains a link to an instruction video for experimenters.



Supplementary Figure 3A: Head-cast design modified to accommodate hair

a) Virtual head model of a subject with grooves to accommodate hair volume. The MRI scan is insensitive to hair so protrusions are added to the scalp extraction to create space. **b)** Photograph of head-cast designed to accommodate hair.

Chapter 4

Experiment 3: Hippocampal theta activity can be detected in MEG during spatial memory with head-casts

Precis

Experiment 1 demonstrated that detection of hippocampal sources in MEG is possible with improved modelling and depends on minimization of co-registration error, while Experiment 2 demonstrated how this minimization can be achieved using flexible head-casts. In this study, the anatomical modelling and head-cast methods are combined to investigate the empirical predictions made in Experiment 1. This study thereby addresses the question of whether it is possible to reliably detect hippocampal activity empirically when the acquisition protocol is optimised (to the best of our ability), and the individual hippocampal surface manifold is included in the generative model. If this proves possible, it will be a step towards the exciting prospect of extending MEG to be a non-invasive, temporally resolved neuroimaging tool for investigation of the dynamics of human hippocampus. One cognitive function of particular interest is spatial navigation which provides a fruitful starting point and a substantial challenge as many of its neural properties have been studied in rodents, uncovering a wealth of predictions and open questions. The focus of this study the methodologies which may pave the way for such possibilities; how the previously presented methods can be combined and to what extent it is possible to validate simulation results using empirical data.

Introduction

Humans display remarkable cognitive skills when navigating through the environment. Like other animals, our behaviour is based on explicit representations of space that must be encoded, stored, and flexibly interrogated. The temporal dynamics of the hippocampal computations which underpin these processes are being characterised in humans using intracranial recordings in epileptic patients. However, studying these using a non-invasive neuroimaging tool such as MEG would be a more generalizable, efficient and unbiased way of exploring these correlates. This prospect has recently been shown to be realistic for several reasons. First, it has been demonstrated using concurrent intracranial hippocampal electrodes (in preoperative epileptic patients) and MEG, that there is a large zero-lag component of the ongoing theta rhythm which is detected by the sensors (Dalal et al., 2013b), and that this can be localised using beamformers (Crespo-García et al., 2016). Second, the electrophysiology and anatomy of the hippocampus lends itself well to MEG; the hippocampus contains a

pyramidal cell layer similar to neocortex (Attal and Schwartz, 2013; Spruston, 2008), it is known to exhibit oscillatory activity in humans (Ekstrom et al., 2005; Lega et al., 2012; Watrous et al., 2011), and information about the environment is encoded across large populations of neurons (Buzsáki and Moser, 2013; Hebb, 1949; Pouget and Driver, 2000). Third, the spatial priors (expected locations of activity) are strong given that cognitive experiments can be adapted from MRI studies with robust hippocampal responses. However, the real-time neural dynamics of these cognitive processes, while extensively studied in rodents and shown here to possess a wide range of quantifiable relationships to behaviour, have received relative little attention in humans (Jacobs, 2014; Riggs et al., 2009; Zhang and Jacobs, 2015). Thus, it is unclear to what extent the neural circuits and their oscillatory properties are shared across species, and whether rodent findings can be successfully translated into human neuroscience and begin to explain, in physical and computational terms at a systems-level, the richness and complexity of our experiences of cognitive feats such as navigation. Crucially also, the cognitive capacities of humans which extend beyond those of rodents may hold answers to more general and clinically useful research questions.

Several functional features of the human hippocampal rhythms have previously been demonstrated. However, while findings based on invasive electrophysiological measurements are informative (but rare), they are not necessarily representative of the general population, or healthy brains. On the other hand, they do rely on a known ground truth (i.e. that the measured signal is hippocampal), whereas the same is not true for MEG studies. Nonetheless, theta oscillations (~4-8Hz) during encoding and retrieval of spatial information have been found in MEG by several research groups, and these studies suggest (but do not show) that the signal is hippocampal (Backus et al., 2016; Cornwell et al., 2012; Dalal et al., 2013c; Jacobs et al., 2013; Kaplan et al., 2014, 2012b; Lega et al., 2012).

Theta oscillations and changes to them in relation to behaviour have a wide range of interesting and informative properties. For example, theta power has been shown to selectively correlate with retrieval of spatial information that is relevant for navigation (de Araújo et al., 2002). At a circuitry level, theta oscillations modulate the activity of hippocampal place cells (O'Keefe and Recce, 1993) and carry information about spatial location in the oscillatory phase, the latter of which is thought to be used by entorhinal grid cells to compute an animal's location in space (Burgess et al., 2007).

While it has been shown that anatomically, the human hippocampus plays a role in detection of novelty (Kumaran and Maguire, 2007), medial temporal lobe (potentially hippocampal) theta power during encoding of spatial information has been shown to relate specifically to environmental, but not content, novelty (Kaplan et al., 2012b). However, it is not clear whether the human hippocampus analogue of rodent theta is at the frequency range thus far assumed (~4-8 Hz) (Lega et al., 2012), raising questions and perhaps concerns about the conclusions drawn from earlier studies.

In addition to theta, high-frequency gamma power increases in hippocampal activity have also been shown to relate to successful memory encoding and retrieval (Burke et al., 2014; Hanslmayr et al., 2016; Staresina et al., 2016). This suggests that power changes in other frequencies, such as gamma, have an important but not-yet-understood role in hippocampal-based memory functions such as pattern completion.

As MEG has a very high temporal resolution and is minimally susceptible to signal attenuation by tissue or skull surrounding the brain, it is well-suited for measuring dynamic brain activity such as theta oscillations. Despite a large and growing body of modelling and empirical studies documenting the feasibility of MEG for detecting hippocampus (Attal et al., 2007; Backus et al., 2016; Guitart-Masip et al., 2013; Kaplan et al., 2014, 2012b; Riggs et al., 2009), as well as simultaneous invasive and MEG recordings directly demonstrating this claim (Crespo-García et al., 2016; Dalal et al., 2013a), controversy regarding the reliability of these claims persist (Mikuni et al., 1997; Riggs et al., 2009; Stephen et al., 2005) and deeper sources are often omitted from analysis de facto.

However, there are now several methodological advances which directly facilitate examination of hippocampal signals using MEG. Across a range of assumptions about the relationship between brain activity and MEG signals, we found that minimizing co-registration error is the single most important factor in being able to reliably detect hippocampal activity (as demonstrated in Chapter 2). Using a new generation of flexible head-casts for MEG (Chapter 3), it is now possible to meet the co-registration error criteria identified in simulations. The key argument is that accurate information about the anatomy of the brain in relation to the MEG sensors enables reliable assessment of hippocampal involvement because generative models with and without the hippocampus, or with variations in the hippocampal portion, can be compared. This comparison can then be used to probabilistically assess whether the hippocampal portion of the model contributes to a parsimonious explanation of variance and generalizability, or not. Importantly, these head-casts are now compatible with visual

stimulus presentation. Using a well-validated virtual reality (VR) paradigm which is known to engage hippocampus in fMRI (Doeller et al., 2008), MEG (likely hippocampal, tentatively medial temporal lobe) (Kaplan et al., 2014, 2012b), and intracranial EEG (iEEG) (Bush et al., in preparation), we hypothesized that it would be possible to find evidence for hippocampal activity in real data. Thus, this experiment uses a combination of novel and well-validated methods to try to demonstrate the face validity of this approach.

Hypothesis

Hypothesis 1: If we combine an acquisition technique which is optimal for obtaining high SNR data (Experiment 2, head-casts), a well-validated spatial memory task which is known to engage the hippocampus, and explicit source modelling of the hippocampus (Experiment 1), we can detect hippocampal sources in real MEG data.

Hypothesis 2: If this combination of tools is effective, then changes to the hippocampal portion of the generative model should give rise to decreases in model generalizability/fitness (which can be quantified by two orthogonal metrics; Free energy and cross-validation error). Specifically, we predict that if the subject-specific generative model of the hippocampus is correct, then laterally rotating it should decrease the model evidence and increase the cross validation error.

Methods

Participants

We recruited 13 participants (11 men, 2 women, average age = 29.5, SD = 7.4, all right-handed). All subjects gave informed consent and were compensated for their participation (with the exception of one collaborator and one author). The head-cast creation, cognitive task protocols, and MEG scanning were approved by the UCL Research Ethics Committee. All subjects had normal or corrected-to-normal vision, no history of psychiatric or neurological disease or claustrophobia.

Task Design and Structure

During scanning participants were required to accurately encode and later remember the locations of objects placed in a virtual arena within which they could freely move (**Figure 4.1**). The circular arena was surrounded by 3-4 distal cues to be used for navigation. The task consisted of two phases; encoding and test. Both were carried out while the subject was being scanned in an MEG scanner while wearing a custom-built flexible head-cast (see Chapter 3 for methods relating to this).

The experiment was designed such that each subject carried out one pre-scanning familiarization trial and four regular trials (encoding plus test, carried out during scanning) on each of two separate visits to the lab. Each encoding trial consisted of presentation of six separate three-dimensional objects, each in a fixed location within the arena. Only one object was visible at a time. The subjects were instructed to “pick up” objects by moving to the location of the object which caused it to disappear. Immediately following this, another object would appear in its respective location. Following the encoding phase, subjects were presented with a fixation cross followed by one of the six objects where fixation and cue periods were 3s each. During the cue period, the subjects were instructed to construct a vivid mental image of where they saw the object and try to incorporate as much detail into this image as possible. Following the cue period, subject were placed in a random location and orientation in the arena, and instructed to navigate to where they think the object was located and indicate their response by pressing a button, after which a new fixation period would begin. As with encoding, subjects performed 24 test runs in pseudorandom order; four for each of the six objects presented during encoding. During test runs, no objects were visible in the arena (the distal cues were still visible however).

Each subject performed four blocks of 24 trials per visit (plus one familiarization trial on each of the two visits). To control and orthogonalize environmental and object novelty, each trial was either performed in a new virtual reality environment. The object sets and virtual arenas were staggered such that on each new trial, only one was novel (to enable comparison of familiar versus novel objects). Object sets, object order, environment order, and starting locations were pseudorandomised and fully counter-balanced across participants. The experiment was self-paced with the exception of the fixation and object presentation periods. Subjects were given breaks between trials to exit the dewar and remove the head-cast if they wished.

Subjects controlled their movements through the environment with two 2-button MEG-compatible control pads. The four buttons were configured to allow the subject to move

left, forwards, right, and to indicate their response (i.e. location they thought the object being tested was located in).

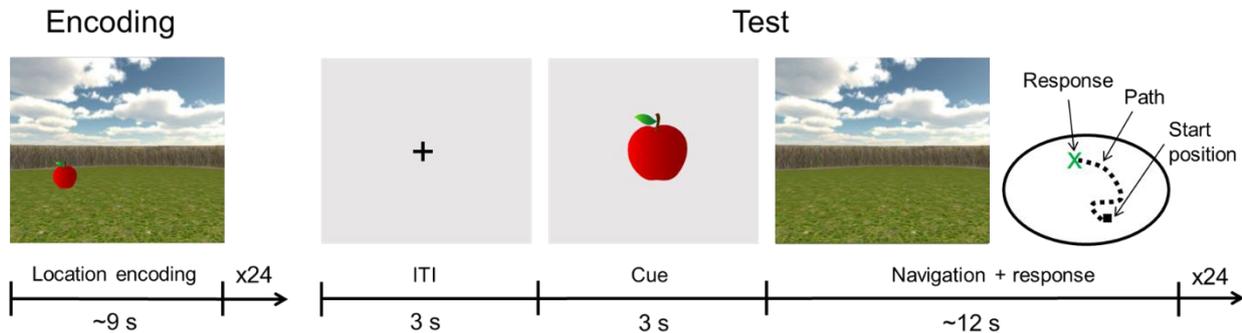


Figure 4.1: Virtual Reality environment and trial structure.

Left: Encoding run showing sample virtual reality environment. Subjects were instructed to encode the locations of the object (chair, duck, helmet, bucket, etc) placed in the environment and to pick it up by navigating to the location. Each of six objects was presented four times with only one object present at a time. Encoding was self-paced and subjects were instructed to focus on remembering the object locations. Right: Test run showing cued retrieval of object location. Each consisted of a 3s inter-trial interval (ITI) followed by a 3s cue period where one of the objects presented during the encoding phase was shown in the middle of the screen. Subjects were instructed to remember, as vividly as possible, the location of the object presented. Afterwards, the subject was placed at a randomized start position in the environment and instructed to navigate back to the remembered object location and indicate their response. Presentation order, object locations and identities were randomized and counter balanced across subjects. Note that the subject only ever saw the environment from an egocentric (first person) perspective.

Virtual Reality

UnrealEngine2 Runtime software (Epic Games, <https://unity3d.com/>) was used to present a first-person perspective viewpoint 2 metres above the ground. All environments were the same size (18 metres in diameter) and shape (circular), and surrounded by a set of distal landmarks to enable orientation within the arena. These were a stone/marble floor, surrounded by mountains, a grassy plane surrounded by trees, a tiled arena surrounded by mountains, and a metal floor surrounded by buildings and towers. All environments also had a clouds in the background and the sun as a consistent light source which could be used for orientation. Participants

practiced the task in an unrelated virtual environment before performing the experiment. 3D objects were presented at different locations inside the environment. The location in the environment and heading location was recorded every 25ms. During the encoding phase, participants were instructed to “collect” the objects by passing through their location. During the cue phase, the objects were presented as 2D images and participants were instructed to remember and subsequently navigate to the object’s original location before indicating a response using the left button box. To trigger the data, a small box was coded to appear at the top left-hand corner of the screen and change from white to black when subjects transitioned from one trial-state to another (e.g. from ITI to cue period). This enabled us to use a photodiode to record and temporally align trial transitions with respect to the MEG data.

MEG Data Acquisition and Inversion

MEG recordings were made in a magnetically shielded room with a 275-channel Canadian Thin Films (CTF) system with SQUID-based axial gradiometers. Data were digitized continuously at a sampling rate of 600 Hz. Subjects wore customized flexible head-casts during recording (methods described in Chapter 3). Fiducial coils were attached to the head-cast and thereby located at MRI-defined nasion, right and left preauricular sites. The coils were continuously energized throughout the experiment for localization of the head(-cast) with respect to the MEG sensors. No subjects deviated more than 5 mm from their starting position. Subsequently to MEG recording, the data were co-registered to an anatomical MRI image.

The inversion parameters used here were kept as similar to those in Chapter 2 as possible for comparability. Firstly, we used the Empirical Bayes Beamformer (EBB) algorithm which bases the source covariance estimate on the data and is free from pre-specified spatial priors and the bias which may accompany them. We analysed the activity during the 3s cue period and baseline corrected this using the immediately preceding 3s inter-trial interval. We used no spatial dimension reduction to ensure valid comparisons across models (thus we used 274 spatial modes, matching the number of functional sensors). We used 16 temporal modes and constrained the solutions to lie on the meshes provided, with the orientation constrained to be normal to the mesh vertices. A Hanning taper was applied to the time-series but no down-sampling or data averaging was done. A Nolte single shell model was used to model the inner skull boundary (Nolte, 2003). For the theta analysis, the frequency of interest

was set to 4-8 Hz, while for the gamma analysis the frequency of interest was set to 60-90Hz.

MRI Data Acquisition and Surface Extraction

Two MRI images were acquired for each subject; one for constructing the head-cast, and one for creating the anatomical models of the cortex and hippocampus for constraining the inverse solutions of the MEG data. Both were acquired using a Siemens Tim Trio 3T system (Erlangen, Germany). While the main criteria for the former was to minimize distortion of the scalp, skin and face, the main criteria for the latter was maximization of spatial resolution. Thus the acquisition times and parameters differed; for the head-cast MRI, a 12-channel head coil was used without padding, and the acquisition time was 3 min 42 s (see Chapter 3 for remaining MRI and head-cast construction protocol details). Conversely, for the high (0.8mm) resolution MRI images, a standard quantitative multiple parameter mapping (MPM) protocol was used with a 32-channel head coil, padding, and 3 x 7 minutes acquisition times for the (see (Weiskopf et al., 2013) for details). Both MRI images were acquired during approximately two weeks prior to the first MEG recording.

The T1-weighted head-cast image was segmented and used to create a virtual head model for 3D printing using standard MRI segmentation procedures in SPM12 (See Chapter 3 for more details). The T1-weighted MPM image was segmented using FreeSurfer (Fischl, 2012) to extract the cortical and hippocampal surfaces for each subject. Freesurfer-based mesh extraction consists of correcting for intensity variations in the image, removing extracerebral voxels, and segmenting the cortical hemispheres and subcortical structures taking into account variability in the histological composition of these structures. Further, the algorithm self-corrects topological defects. The result is a triangular tessellation of each structure; in this case two hemispheres and two hippocampi per subject. These meshes were then used to constrain the inverse solutions by modelling each structure with the assumption of pyramidal cell-generated signals emerging whereby the source locations and orientations were constrained by the mesh. The locations were modelled by the vertices and the orientations were modelled by the normal orientation to these vertices. The number of hippocampal vertices was ~400 for both hippocampi, making up ~2% of the total vertices in the combined model.

The MPM image was co-registered to the head-cast MRI image and the MEG data was subsequently co-registered to the MPM image.

Pre-processing

Epochs corresponding to fixation and cue period were defined as -1000 to 7000 ms relative to the onset of the fixation cross. We included 1s of padding on either side of the baseline and cue period pairs to avoid analysis-induced artefacts. The data were baseline corrected whereby the mean of the pre-cue (baseline) period was subtracted from the activity during the 3s cue period. The data were then high-pass filtered 0.5 Hz, low-pass filtered at 150 Hz, and the power-line interference was stop-band filtered away by removing 48-52 Hz. Data were analysed with SPM12 (<http://www.fil.ion.ucl.ac.uk/spm/>) within MATLAB 2014a (The MathWorks).

Eye blinks were detected from a frontal channel and detected eyeblinks were used to obtain an average blink time course, based on which a principle component analysis was used to obtain templates of the spatial topography related to blinks. The main component obtain was regressed out of the data before proceeding with analysis. Artifactual epochs were detected by visual inspection and rejected using the FieldTrip visual artefact rejection tool.

Model Comparison

To assess the role of the hippocampus in generating the observed signals, we specified, for each subject individually, a set of forward models which varied with respect to the hippocampal mesh extracted from their MRI image. The forward model constitutes part of the generative model describing how the data arose. All other components of the generative model were left the same across all models tested. We explored here the model evidence (approximated using Free energy, F) as well as the cross validation error (CVE) values associated with each inversion carried out with a different model. These two metrics are independent but it is worth noting that the inverse solutions are optimised with respect to Free energy. Both Free energy and CVE allow formal comparisons of different models of data. While Free energy works in a Bayesian framework and can be conceptualised as a Bayes factor, CVE reflects how well a random subset of sensors can be predicted based on the remaining sensors, given the generative model (see introductory section *Free Energy* and next

section *Cross Validation Error* for more details of how these are calculated). Both values are used here as relative or comparative measures; that is, we compare the Free energy and CVE values from different models against each other. The model comparisons presented here provide a complement to the simulations presented in Chapter 2 in the form of empirical substantiation of the assumptions made, given that the acquisition requirements are met by using head-casts.

Figure 4.2 shows an example of the set of generative models tested against the data recorded. First, a cortical (and hippocampus-free, **a**) model is compared against a combined model which includes the individual subject's hippocampus (**b**). This is the basic model comparison which assesses whether or not modelling the hippocampus at all, facilitates explaining variance in the data. Next, we add different degrees of lateral rotation to each hippocampal mesh (**c-f**).

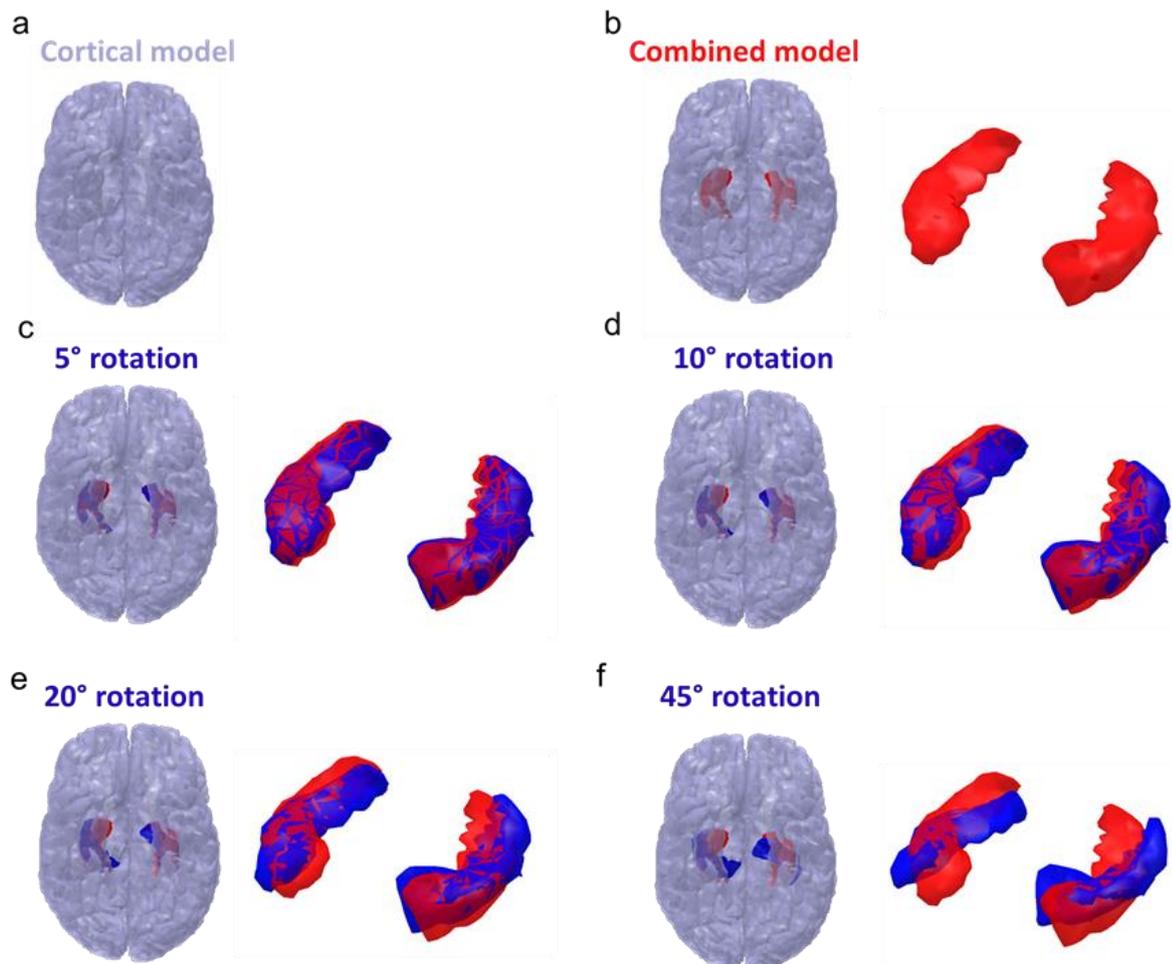


Figure 4.2: Anatomical models; cortical model, combined model, combined models with shifted hippocampi.

a) Cortical model comprising only the cortical surface. The cerebral cortex is used as a model of putative cortical sources. This is the standard model/method used in most non-volumetric MEG inversion algorithms. **b)** Combined model comprising the cortical surface and the hippocampal surfaces. **c)** 5° rotation model. This is equivalent to the combined model but the lateral axis of the hippocampus is rotated by 5° and shown in blue. Note that the red hippocampal mesh is the non-rotated mesh, included for visualisation of the difference. **d)** Same as c) but with 10° of rotation added to the hippocampal meshes. **e)** Same as previous but with 20° of rotation. **f)** Same as previous but with 45° of rotation. The meshes shown here are a representative example taken from a single subject but note that each generative model is subject-specific and based on an anatomical MRI image with 0.8mm resolution.

Cross Validation Error

Cross validation is a model validation method used to assess how well a given model will generalize to an independent measurement. Here, we apply this method to the sensor-level signals and ask how well these can be predicted using different generative models of the data. The validation component of this method consists of leaving out a subset of the data, and measuring how well this subset can be predicted. Thus, we examine the predicted signals in the left-out sensors.

We can then compare the different generative models with respect to the error in these predictions. The units of these errors are femtoTesla (fT). In this set-up, we take out 10% of the sensors (equal to 27 sensors), and use the remaining 90% (247) to create a model. We then use this model to predict the signals observed. In this analysis we then calculate the average cross validation error across four iterations per dataset. The figure below shows the different time-courses measured and predicted at a random sensor over the course of 1 second of the cue period for a single subject (**Figure 4.3**). The black line shows the measured data while the red shows the predicted data when the generative model includes the correct (straight) hippocampus. The blue line shows the predicted data at the sensor when the hippocampi are rotated by 45° (**Figure 4.2f**).

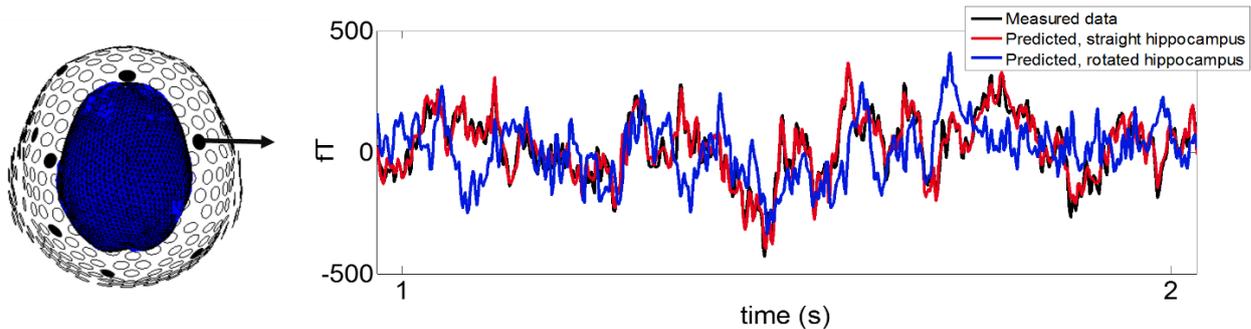


Figure 4.3: Cross validation method

Left image shows a random subset of sensors left out of the analysis and subsequently predicted using generative models containing variations of the hippocampal mesh. The generative models can be compared on the basis of their error in predicting the measured signals across the randomly left out sensors. Right plot shows an example of data measured (black line), predicted with a generative model which includes the correct (straight) hippocampus (red line), and predicted with a generative model where the hippocampal mesh is rotated 45° laterally (blue). For illustration, random sensors were marked and for clarity, only 1s was included in the plot (although all 3s of the cue period were entered into the analysis).

Results

Free Energy

We created a standard cortical and a combined (including the hippocampus) generative model for each subject (**Figure 4.2**) and compared the Free energy values obtained upon inversion of the cue period where subjects were told to remember the spatial location of an object. To first establish that this improved the model evidence, we found a significant (>3 log units, BOR <0.001) increase from inclusion of the hippocampus in all subjects (**Figure 4.4**). Consistent with simulation results (Chapter 2), we observed evidence in favour of hippocampal engagement during the cue period of the spatial memory task. Notably, the Free energy values and thus differences scale with the amount of data, and so are far larger here where we have ~96 trials per subject compared to in simulations which are based on single trials.

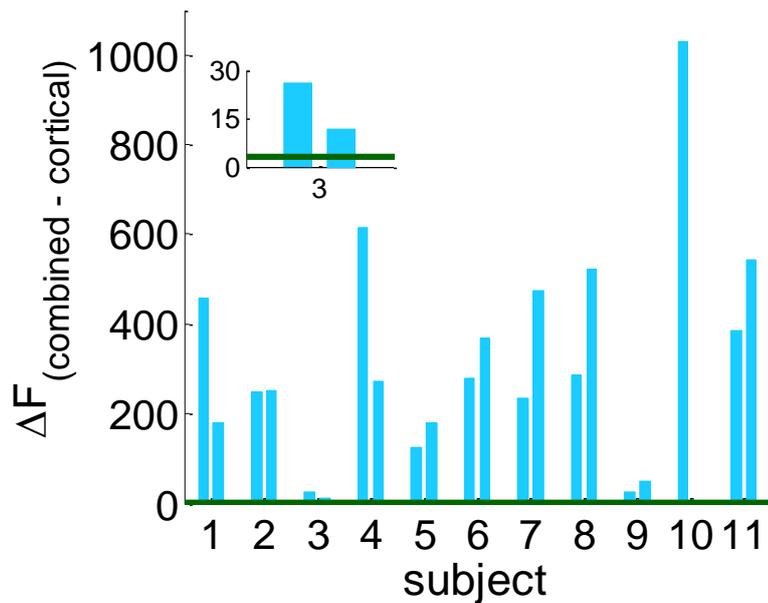


Figure 4.4: Free energy differences: cortical versus combined models

Anatomical model comparison results for 11 subjects. Bars represent Free energy value differences when the anatomical combined and cortical models at 4-8Hz during the 3 s cue period where subjects are instructed to remember the location of the object presented. There are two bars per subject, representing the two separate visits and datasets. Note that subject 10 was only scanned once however. All subjects are (well) above the significance threshold of 3, implying that the advantage of including the hippocampal mesh in the generative model outweighs the increased complexity introduced. Inset shows zoomed-in view of results for subject three; both bars are well above the significance threshold.

Next we were interested in the effect of rotating the hippocampal mesh on Free energy. Here we therefore compared the combined model (in this context the straight model) to models with rotated hippocampi. This analysis revealed that the average model evidence between the model with a straight versus rotated hippocampus is significant in the 4-8Hz band even when this rotation is only 5° (**Figure 4.4a**). This effect is significant at group level across all degrees of rotation (average $\Delta F > 3$). The involvement of hippocampal generators in the recorded MEG signal is thus likely to be well-captured by the generative model created on the basis of subject-specific MRI images. At 45°, 21 of 22 model evidence differences are significant at < -3 log units difference.

An interesting follow-up question is whether this finding generalises across frequency bands. We found that constraining the frequency to the gamma band (60-90 Hz), we

were also able to find decreases in model evidence which scaled with rotation of the hippocampal mesh (**Figure 4.5**). Here, the average model evidence difference reaches significance at 20° of rotation.

Finally, as a control condition, we shuffled the lead field labels such that the relationship between anatomy and sensors was destroyed (**Figure 4.5c**). This shows that there is no bias inherent in the analysis, and that the anatomical information relates to the signal recorded.

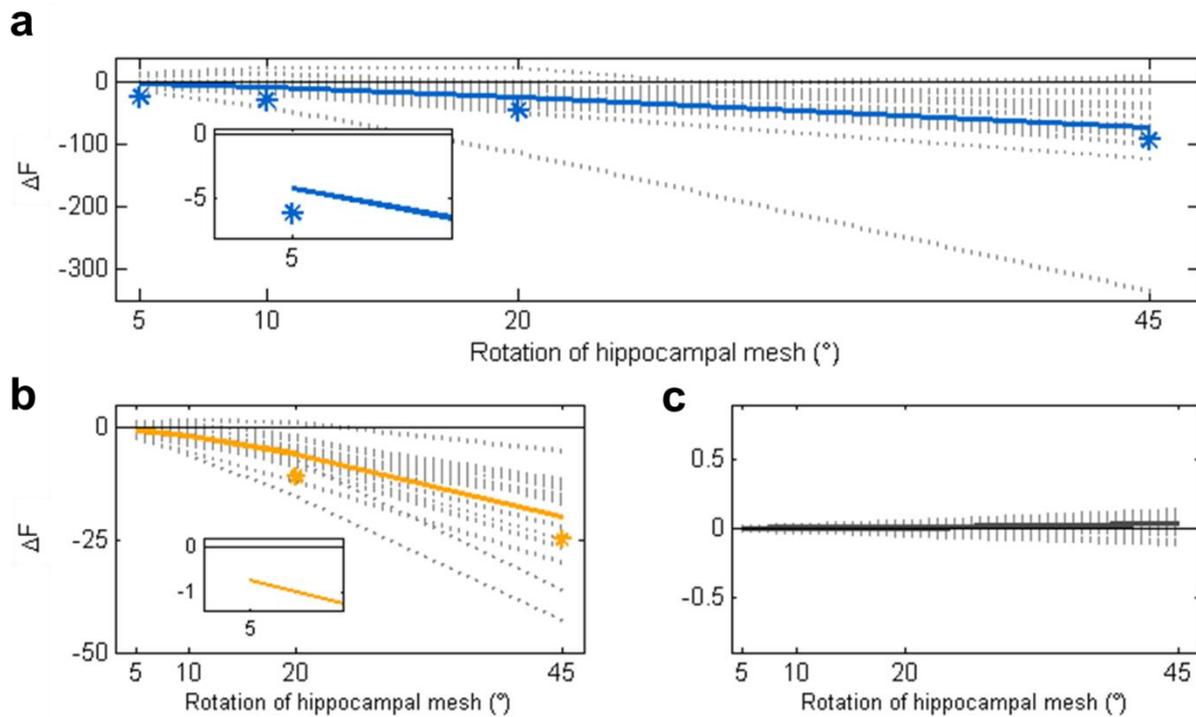


Figure 4.5: Free energy differences: straight versus rotated hippocampi

a) Free energy difference increases as hippocampal mesh rotation increases. Average difference across all subjects shown in blue. All degrees of rotation are significant (mean < -3). Theta (4-8Hz) analysis. We find that comparing the straight to rotated hippocampi (rotated-straight), the change in model evidence scales with degree of rotation of the hippocampal mesh. Inset shows significance of mean model evidence difference level (< -3) at 5° rotation.

b) Same as a) but applied to gamma (60-90Hz) frequencies. Average ΔF becomes significant at 20° of lateral rotation. Average across subjects shown in yellow. Inset shows mean model evidence at 5° of rotation (not significant).

c) Control analysis where lead field sensor labels are shuffled to destroy relationship between anatomy and sensors.

Table 5 shows the mean Free energy differences and accompanying BOR values for the two frequency bands tested as well as the lead field shuffle condition. Values are calculated across all subjects. We find that while the mean Free energy difference reaches significance at 5° rotation in the theta band (with a significant BOR value), this is true at 20° in the gamma band, and never in the shuffled lead field condition.

Table 5: Mean Free energy differences and Bayes Omnibus Risk values across frequency bands/conditions

	Theta		Gamma		Shuffled	
	Mean F	BOR	Mean F	BOR	Mean F	BOR
5° rotation	-4.252	2.1774e-09	-0.618	<0.001	-0.0002	0.875
10° rotation	-10.226	1.8459e-10	-1.774	2.478e-09	<0.001	0.875
20° rotation	-26.268	8.4852e-13	-5.419	4.8853e-16	0.005	0.874
45° rotation	-73.909	3.0238e-14	-18.534	3.4484e-19	0.0327	0.816

Cross Validation Error

As a second independent measure of performance of the different models, we calculated the cross validation error when leaving out 10% of the sensors over 4 folds. This analysis showed approximately the same result as the Free Energy findings (**Figure 4.6**). Table 6 shows the p-values, t-values, and degrees of freedom (df) for each frequency band or condition tested. We find that in the theta band, the effect of mesh rotation is significant at group level at 10° while in the gamma band it is significant at 20° or more. When the lead fields are shuffled as a control condition, there is never a significant effect of mesh rotation.

In both frequency bands tested, subjects consistently showed consistent increases in error across degrees of rotation. In both frequency bands therefore, the activity measured can therefore be attributed to the hippocampus. Together, the Free Energy and cross validation error results support the case that we are measuring hippocampal sources (confirming the second hypothesis stated).

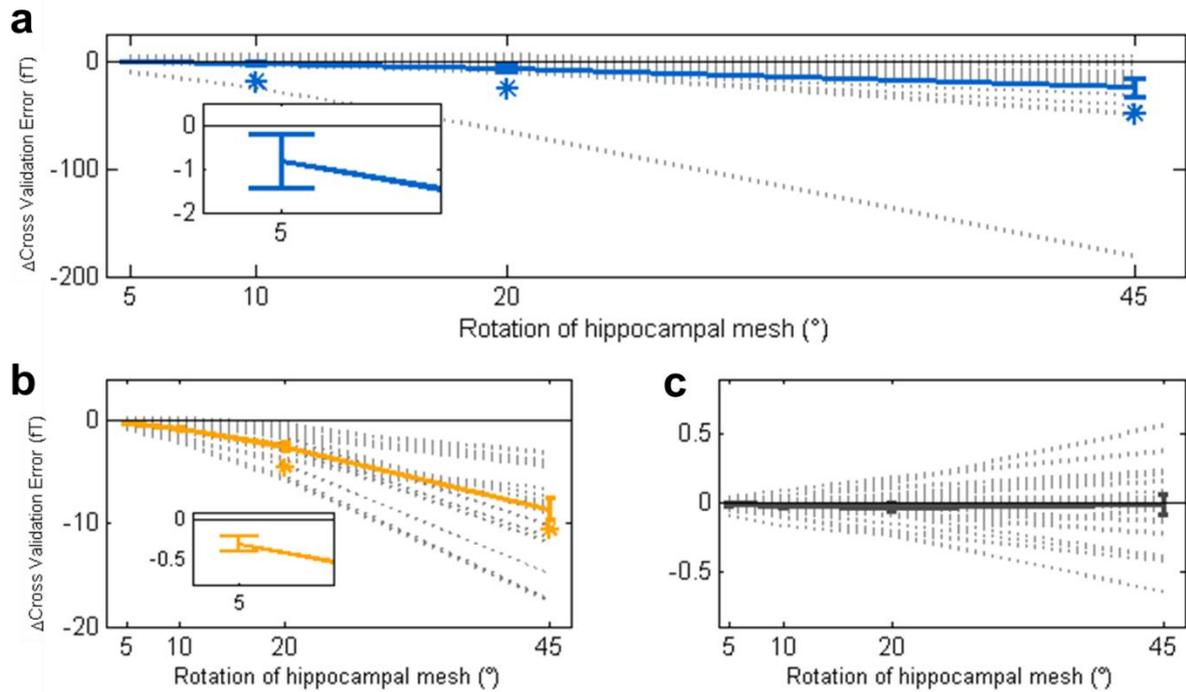


Figure 4.6: Cross validation error: straight versus rotated hippocampi

Same format as Figure 4.5 but reflecting cross validation error changes. **a)** Cross validation error increases as hippocampal mesh rotation increases. Average difference across all subjects shown in blue. All degrees of rotation are significant. Theta (4-8Hz) analysis. We find that comparing the straight to rotated hippocampi (rotated-straight), the change in model evidence scales with degree of rotation of the hippocampal mesh. Inset shows zoomed-in view of 5° rotation data where the mean is below zero. **b)** Same as a) but applied to gamma band data (60-90Hz). Average cross validation error shown in yellow becomes significant at 20° of lateral rotation. **c)** Control analysis where lead field sensor labels are shuffled to destroy relationship between anatomy and sensors. Errorbars reflect SEM.

Table 6: Cross validation statistics for mesh rotation analysis

	Theta			Gamma			Shuffled		
	p-val	t-val	df	p-val	t-val	df	p-val	t-val	df
5° rotation	0.094	-1.370	18	0.834	0.996	17	0.066	-1.579	18
10° rotation	0.046	-1.778	18	0.832	0.988	17	0.077	-1.487	18
20° rotation	0.018	-2.267	18	3.426e-07	-7.434	17	0.120	-1.214	18
45° rotation	0.007	-2.703	18	4.781e-08	8.539	17	0.418	-0.210	18

Due to a technical problem with the inversion in the gamma band for a single dataset, the data was not included and the degrees of freedom equals 17 and not 18 as with the others.

Discussion

This experiment was concerned with the empirical validation of the hippocampal source modelling method presented in Chapter 2. We used flexible and subject-specific headcasts to minimize co-registration error and head movement, and MRI-based anatomical modelling of the hippocampus for explicit evaluation of hippocampal engagement under optimal (low co-registration error) recording conditions. We hereby demonstrate empirically the contribution of hippocampus-specific activity to the measured signal, validating the modelling and simulation results presented in Chapter 2. Through application of Bayesian model comparison and cross-validation, we found that lateral rotations of the hippocampal portion of the generative model significantly decreases the predictive power of the model as a whole (quantified using two independent metrics), even when these errors are as low as 5° .

We can also confirm our second hypothesis (sensitivity to subtle distortions of the hippocampal portion of the generative model). We found evidence for this across two independent metrics; Free energy and cross validation error. We found that for theta band activity, we are sensitive to lateral rotations of the hippocampi as small as 5° with either method. We find this to be significant at the group level. With Free energy, we take significance as an average difference between the two models of 3 or more (where one model is 20 times more likely than the other, whereas for cross validation error we assess significance using a one-tailed t-test. It is important to note that we only did 4 folds of cross validation, which is less than what is commonly used. However, the results should in theory strengthen with more iterations.

It is important to note here that with rotated hippocampi, there is an inherent advantage from a modelling perspective; as the dipole orientations are rotated away from those of the cortical mesh in the medial temporal lobe, the hippocampal mesh becomes an increasingly better model for explaining noise or artefacts. Thus, the rotation should in theory not only be neutral with respect to the ability of the model to explain data, but actually increase it if the source was non-hippocampal.

From a methodological perspective, another set of exciting possibilities remain. For example, it would be interesting to assess the specificity of the anatomical structures with respect to individual variation; if one were to swap the hippocampi across subjects, would it be possible to recover the true anatomy using model comparisons?

I.e. can this data and analysis approach also be used to obtain structural information from the MEG data?

Several open questions remain to be addressed in this dataset. These include questions regarding the relationship between theta and behavioural performance (Crespo-García et al., 2016; Ekstrom et al., 2003; Watrous et al., 2011), the significance of theta during encoding (Jensen and Lisman, 2005; Lega et al., 2012), and movement-related changes (Kaplan et al., 2012a). Finally, in terms of pertinent cognitive questions, theta power increases in hippocampus which are coupled to theta in medial prefrontal regions could be explored using this dataset (Backus et al., 2016; Guitart-Masip et al., 2013; Kaplan et al., 2014). Further, several very close-by structures and cell types are also known to be modulated by, and thus exhibit theta oscillations during navigation. These include entorhinal grid cells, subicular place and direction cells, and cells of the parahippocampal cortex (Burgess, 2008; Cornwell et al., 2008). In the case of grid cells, an interesting prediction would be that such modulations, in either or both theta and gamma ranges, would likely follow six-fold symmetrical spatial patterns which represent space.

Regarding the different cell types and structures involved in coding space, this poses a challenge as they comprise a large number of potential sources. Moreover, it is likely that these structures or a subset hereof are concurrently active with the hippocampus, but unclear whether or how they give rise to a measurable MEG signal. It follows that it would be interesting to model these structures and evaluate their contribution to the measured signal explicitly within the Bayesian and/or cross validation framework. However, doing so may be non-trivial, as the pyramidal cells are generally not distributed or oriented in a laminar fashion, raising doubt about how most appropriately to model them with respect to the MEG signal.

In summary, the findings accord with the results found in simulations; the addition of a hippocampal mesh increases the model evidence (**Figure 4.4**), and this improvement depends on the correct orientation of the hippocampal mesh (**Figure 4.5,6**). Specifically, we have shown that we are sensitive to minute distortions in the generative model of hippocampus, empirically validating the simulation approach presented in Chapter 2, and answering the question raised at the beginning of this thesis (is it possible to detect hippocampus using MEG?) with a 'yes'.

Key points

- Lateral rotations of the hippocampal portion of the generative model significantly decrease its predictive power, even when these errors are as low as 5°.
- Hippocampus-specific activity can be recorded with MEG using head-casts and inferred using appropriate source modelling.
- Both hippocampal theta (4-8Hz) and gamma (60-90Hz) activity can be imaged using this combination of methods.

This chapter derives in part from papers 3 and 4: “Using head-casts to image hippocampus with MEG” Sofie S Meyer, Daniel Bush, James A Bisby, Aidan Horner, Neil Burgess and Gareth Barnes (in preparation), and “MEG sensitivity to hippocampal dipole orientations” Sofie S Meyer, James Bonaiuto, Daniel Bush, James A Bisby, Aidan Horner, Neil Burgess and Gareth Barnes (in preparation).

Chapter 5

Experiment 4: Optimal configuration of optically pumped magnetometers for detecting hippocampal signals

Precis

Experiments 1, 2 and 3 showed that minimizing co-registration error and head movement are key to hippocampal detectability, that head-casts can be used to minimize both while maintaining comfort, and that by using head-casts empirically, successful hippocampal detection is possible. Nonetheless, the hippocampus is still a deep brain structure and therefore inherently difficult to measure activity from because the intensity of magnetic fields fall off with the square of distance from the source. The source-scalp distance thus poses a constraint on the SNR achievable. However, a second and potentially larger SNR constraint comes from the several centimetres of scalp-sensor distance with a normal MEG scanner. This distance is large due to the requirement for cryogenic cooling of the sensors, and in part also the “one size fits all” design. Therefore, if we want to detect hippocampal signals with higher SNR, reducing the scalp-sensor distance is a good starting point.

This has recently been made possible as optically pumped MEG sensors which do not require cryogenic cooling have recently become commercially available. These sensors can be placed directly and geometrically flexibly on the scalp, thus greatly reducing the scalp-sensor distance and removing the “one size fits all” problem (as arrays can be optimised on a subject-by-subject basis). However, the flexibility of placement also changes the nature of the modelling uncertainty: although secure placement of the sensors on the head removes head movement errors, co-registration errors can originate from errors in either the orientation or location of sensors, be these with respect to the head or other sensors.

Thus, in order to understand the potential benefits as well as new sources of uncertainty better, it would be useful to carry out a set of simulations where activity is known to be hippocampal, and the detectability improvements and pitfalls can be directly quantified. In this chapter, the geometrical flexibility of these new sensors is thereby dealt with, both in terms of the different kinds of modelling error it potentially introduces, the effects of these, and how one might use the flexibility to spatially configure an array of sensors for detection of hippocampal activity.

Introduction

At present, MEG experiments are carried out using an array of superconducting coils, each coupled to a superconducting quantum interference device (SQUID). The operation of these sensors relies on cryogenic cooling, typically liquid helium which

not only makes the scanners costly to acquire and maintain, but also imposes a limitation on the minimal separation between the scalp and sensors of several centimetres. This distance is non-trivial given that magnetic field strength decreases with the square of distance from the source.

Over the past decade, rapid progress in quantum technology and engineering has led to the development of small, non-cryogenic magnetometers which detect magnetic fields with the same sensitivity as SQUIDs (Shah and Wakai 2013). These Optically Pumped Magnetometers (OPMs) operate at $\sim 150^{\circ}\text{C}$ yet can be constructed such that the sensitive volume is only a few millimetres from the surface which remains at approximately room temperature. This is crucial, as it enables placement of the sensors directly on the scalp. OPMs thus offer drastic improvements in the sensitivity of measurements because of the decreased distance between source and sensor, as well as the potential for flexibly arranging the sensors according to brain regions of interest.

As OPMs can be placed on the scalp surface, the largest improvement over SQUID systems will be at the cortical surface where the source-sensor distances will undergo the greatest reduction relative to a SQUID set-up. Empirical demonstrations of OPMs for different superficial sources have recently been shown: evoked responses can be detected from both auditory (Johnson et al., 2010, 2013; Xia et al., 2006) and somatosensory stimulation (Johnson et al., 2010; Sander et al., 2012). Critically, recent studies have also demonstrated the feasibility of multi-channel arrays (Johnson et al., 2013; Kim et al., 2014) and improvement of the sensitivity to match that of SQUIDs at $\sim 10 \text{ fT}/\sqrt{\text{Hz}}$ (Shah and Wakai, 2013; Tiporlini and Alameh, 2013). As sources within the brain get deeper however, the relative 8-10 fold sensitivity gain at the cortical surface decreases to around a factor of 1-4 for deeper brain structures in beamformer-based simulations (Boto et al., 2016). Whether a more structure-specific approach using spatial priors to test for deep activity will yield similar or better results is relevant for understanding the potential of these sensors but currently unclear. This chapter is concerned with evaluating the expected sensitivity improvements when using OPMs to detect hippocampal sources. To examine this, anatomical model comparison for hippocampal sources (as described in Chapter 2) is used. The rationale behind this approach is that including a hippocampal mesh should give a more parsimonious generative model which will return a higher model evidence value if sources are hippocampal.

Interestingly, as the OPM technology does not require a fixed array of sensors (the equivalent of an MEG dewar), we become prey to new and different forms of potential modelling errors, as we no longer necessarily know the positions and orientations of the sensors relative to one another or in absolute terms. In this chapter, the advantages of using an OPM system to localize electrical activity in the hippocampus is assessed in relation to this uncertainty: we ask how both independent and systematic location and orientation errors affect our ability to reliably detect hippocampal sources through model comparison-based inference. Finally, we conclude with how to potentially minimize these errors in an empirical context (see also General Discussion).

Hypothesis and objectives

If OPM sensors have the same noise sensitivity as SQUIDs but can be placed directly on the scalp, then they should be able to detect weaker signals simply by virtue of being closer to the sources generating them. Here we are interested in the extent to which this is true for hippocampal signals.

If we can manipulate four different kinds of potential modelling errors independently (systematic and independent errors in both sensors location and orientation), then we can identify how detrimental these are to our ability to reliably detect hippocampal activity. This in turn can be used to inform the design of an OPM head-cast.

Methods

Simulation set-up

As in Chapter 2, the simulation and reconstruction pipeline consisted of three main steps: 1) simulation of a single hippocampal dipole patch, 2) reconstruction hereof with two generative models, one with a nested hippocampal mesh and one without, and 3) model comparison of the two models through comparison of model evidence.

The simulation parameters were similar to those in Chapter 2: a sinusoidal waveform of 40 Hz was simulated for 500 ms as a dipole with FWHM of ~6 mm, oriented perpendicularly to a randomly chosen vertex in the hippocampal mesh. Instead of SNR, here we varied the total effective dipole amplitude to assess hippocampal detectability at varying source strength. $10 \text{ ft}/\sqrt{\text{Hz}}$ Gaussian white noise was added to the data at sensor-level. At each source strength, 30 datasets were simulated and the average model evidence difference across the corresponding model comparisons was

computed. As in Chapter 2, we used only one hemisphere and one hippocampal mesh for simplicity.

For the source reconstruction, we used the Empirical Bayes Beamformer (EBB) which does not require explicit specification of anatomical priors as it makes use of the entire source space (i.e. the full generative model, whether or not it includes a hippocampal mesh). We thereby avoid the problem of having to specify a ratio of hippocampal to cortical priors, and of deciding whether to specify the correct (simulated) hippocampal priors, as was done with the Multiple Sparse Priors (MSP) algorithm in Chapter 2. Other parameters were the same as in Chapter 2 for comparison, with the exception of the number of spatial modes which was set to 270, meaning that we did not perform any spatial dimension reduction as the number of sensors simulated was also 270 for each sensor type. The time and frequency windows were set to match those simulated such that these were 0-500 ms and 0-80 Hz. We applied a Hanning taper to the time series but did not do any down-sampling or data averaging.

To evaluate the effect of moving sensors closer to the scalp, we constructed two virtual sensor arrays. These were both based on the MRI-derived scalp extraction for a single subject (the same subject's brain used to construct the generative models). To explore the flexibility of positioning of OPMs and maintain comparability with SQUIDs, we scattered 270 sensors at random locations but specific distances from this scalp surface. To simulate an OPM array, this distance from the sensitive volume to the scalp was 3 mm (**Figure 5.1a**) and to simulate a SQUID array it was 3 cm. The noise floor of the two sensor types was assumed to be equal ($10 \text{ fT}/\sqrt{\text{Hz}}$) and both were configured as radial magnetometers, i.e. detecting one radial measurement each (although OPMs can in practice be set to record both radial and tangential fields). Thus, the OPM and SQUID arrays differ only in their distance from the scalp.

In this simulation set-up, the distance from the centroid of the hippocampal mesh to the closest sensor was 8.44 cm for the SQUID system and 5.81 cm for the OPM system. Likewise, the average distance from the hippocampal centroid (across all 270 sensors) was 12.81 versus 10.26 cm for the SQUID and OPM systems respectively.

Co-registration errors

We used two different forms of co-registration error to investigate the tolerance of OPMs to errors. We first simulated systematic shifts in sensor locations where sensors

stayed constant with respect to each other, but changed relative to the head location (**Figure 5.1b**). This is similar to standard co-registration error in (standard SQUID-based) MEG where the head may move relative to the sensors, or the co-registration between anatomical and functional data may be shifted due to mislocalization of one or a subset of fiducial coils. Due to the OPM-specific flexibility of sensor positioning, we also simulated independent errors in sensor location where sensors were shifted relative to one another, breaking the array's geometry (**Figure 5.1c**) as well as the spatial relationships of the sensors and the head.

For comparability across sensor arrays, each dataset was inverted the same two generative models (with versus without hippocampus) with both sensor arrays (OPMs and SQUIDs). This enabled direct comparison of the performance of the two different sensor arrays by comparing the point at which the hippocampus became discernible. We can thereby identify the parameters (source strength and sensor distance) under which it is possible to reliably detect the hippocampal signal through model comparison.

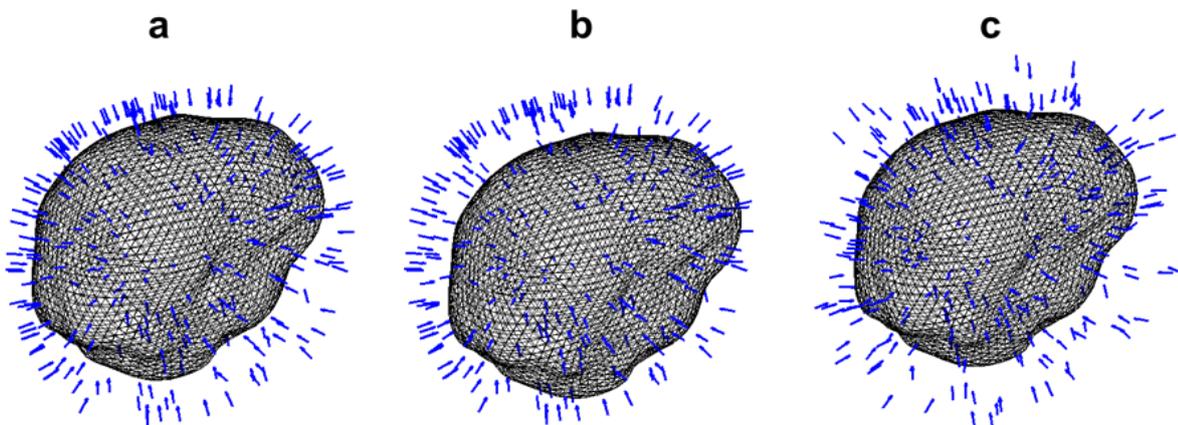


Figure 5.1 Schematic showing the different types of sensor location error for OPM arrays.

a) Standard locations of OPM sensors 3 mm from the scalp surface. b) Systematic shifts added to the sensor locations. Sensors are shifted together in a random direction but maintain the spatial relationships to one another (equivalently to standard co-registration error where the location of the head relative to the fixed sensor array is uncertain). c) Independent shifts added to the sensor locations. Here the location of a given sensor is randomly perturbed which breaks the spatial relationships between sensors, as well as those between the sensors and the brain. Here error types are shown for simulated OPM sensors (3 mm stand-off from the scalp) but the same types of shifts were added to the simulated SQUID sensors (30 mm stand-off from

the scalp). The same principle was used for the sensor orientation error simulations. Blue unit vector arrows are used for illustration purposes only and do not represent sensor size or shape.

Results

Effect of sensor stand-off distance

We first investigated how large of an improvement in detectability of hippocampal sources could be expected from moving the sensors closer to the scalp. To address this, we compared generative models with and without the hippocampus across a range of source strengths. As expected, we found that smaller scalp-sensor distances equates to greater sensitivity (**Figure 5.2**): While OPMs (blue) give model evidence differences of 3 (where the combined model is 20 times more likely) at ~10 nAm signals, SQUIDs only do so at ~50 nAm (**Figure 5.2B**). Notably however, the significance depends also on the Bayes Omnibus Risk values (Table 7) which show that while OPMs reach significance at 20 nAm, SQUIDs do so at ~100. Thus we are approximately 5 times more sensitive to hippocampal sources using OPMs.

Table 7 shows the BOR values for the OPM and SQUID model comparisons across source strengths. While the OPM results reach significance at 20 nAm (but not at 100 nAm), SQUID based simulations reach it at 100 nAm.

Table 7: Bayes Omnibus Risk values for OPM and SQUID simulations across source strengths

BOR	1 nAm	5 nAm	10 nAm	20 nAm	50 nAm	100 nAm	200 nAm
3 mm	0.711	0.377	0.110	0.043	0.034	0.166	0.035
30 mm	0.711	0.689	0.552	0.320	0.081	0.044	0.038

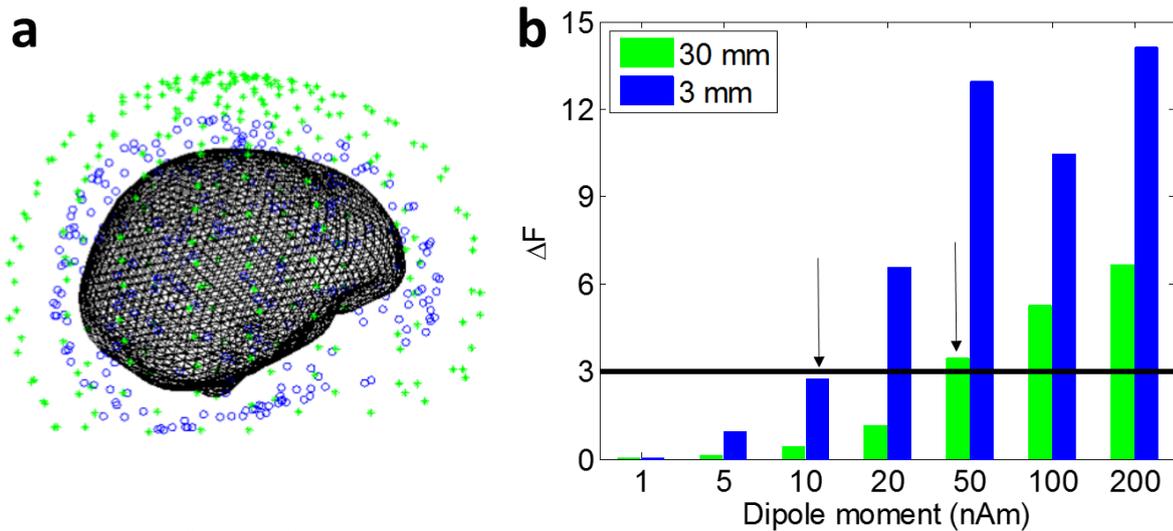


Figure 5.2 ΔF values across simulated source strengths for OPM and SQUID arrays.

A) Depiction of sensor locations with OPM sensors virtually placed 3 mm from the surface of the scalp (blue) and SQUID sensors virtually placed 30 mm from the scalp (green). **B)** Bar plot showing model comparison results for simulated hippocampal sources reconstructed with either an array of OPM sensors (blue) or SQUID sensors (green). Each bar encodes the average model evidence difference across 30 simulated hippocampal datasets reconstructed with a generative model which includes the hippocampus, and one which does not; $\Delta F_{\text{anatomical}} = F_{\text{combined}} - F_{\text{cortical}}$. The black line marks the significance threshold ($\Delta F=3$) where the combined model is 20 times more likely than the cortical. Bringing the sensors 2.7 cm closer to the scalp equates to increasing the sensitivity to hippocampal signals approximately 5 fold (~20 versus ~100 nAm strength required for detection, see Table 7).

Effects of independent and systematic errors in sensor locations

In the previous analysis we assumed that the forward model is accurate. In reality however, there are several factors which may impede the modelling accuracy, such uncertainty about the location and/or orientation of the sensors with respect to each other, and/or with respect to the head. We first investigated whether and to what extent sensor location errors obstruct our ability to discriminate between generative models and thus detect hippocampal sources.

These results demonstrate that errors do negatively affect hippocampal detectability (**Figure 5.3**). Specifically, too much independent sensor location error (e.g. 10 mm, orange lines) completely eliminates hippocampal detectability with both OPMs and SQUIDs (**Figure 5.3a,b**). With smaller errors, increased source strength can to an

extent compensate for errors (compare green and blue lines, **Figure 5.3a**). Similarly increasing the SNR by using OPMs affords improved tolerance of errors. For example, when there is 5 mm of independent error added to the sensor locations (green lines), it is only possible to detect hippocampal sources if the sensors are close to the head (**Figure 5.3a,b**). Even with very strong hippocampal sources (200 nAm), the average SQUID model evidence differences does not reach significance. Results are similar but slightly better when adding systematic errors (dotted lines, **Figure 5.3c,d**). With 10 mm of independent error (dotted orange lines), it is generally possible to detect hippocampal sources above 20 nAm with OPMs but not SQUIDs, regardless of source strength. Although there seems to be an effect of error type (solid versus dotted lines in **Figure 5.3c,d**), this is not significant (Bayes Omnibus Risk comparing systematic and independent errors is 0.077 for OPMs and 0.105 for SQUIDs).

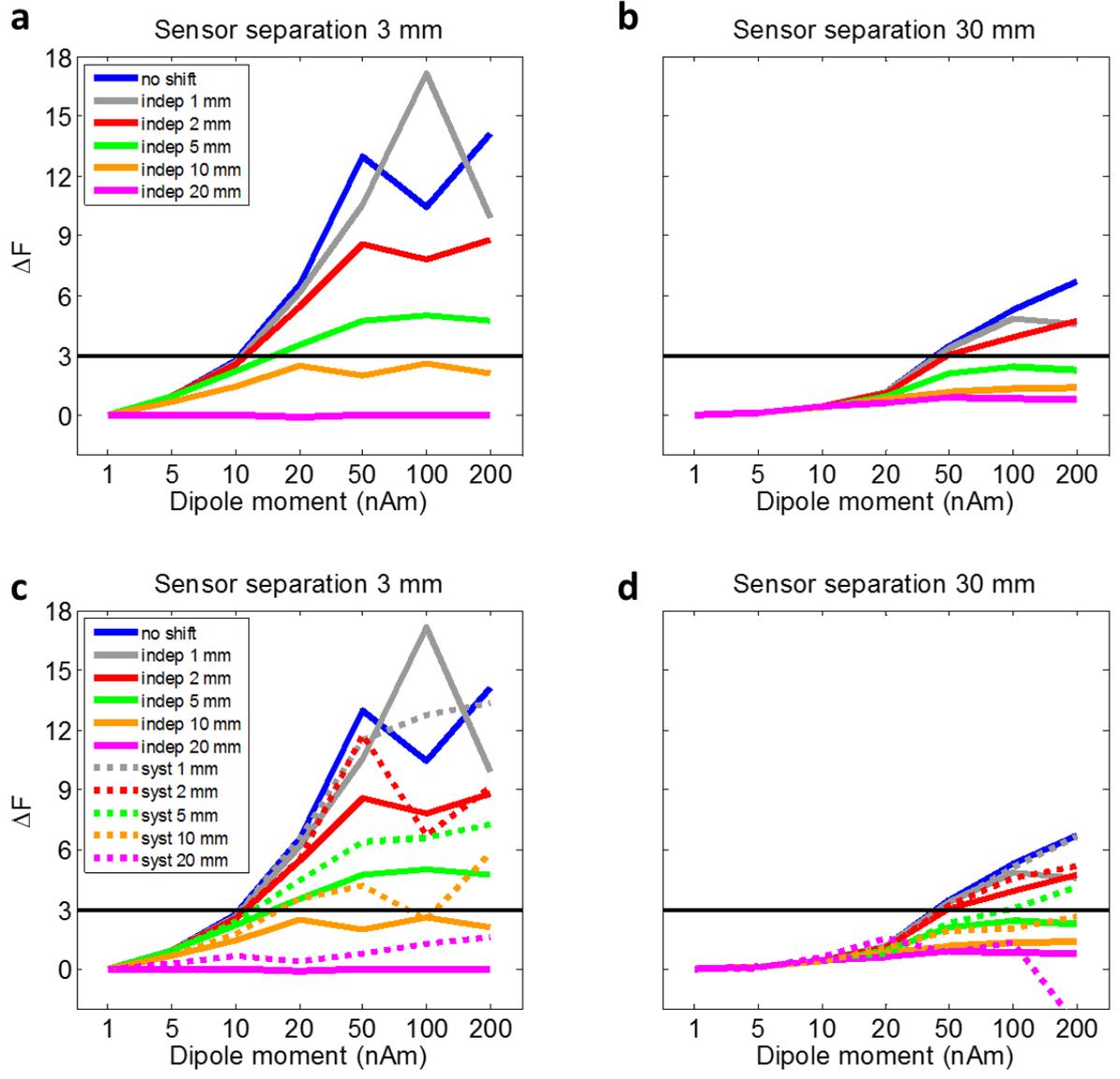


Figure 5.3 Effects of systematic and independent sensor location errors on ΔF across simulation strengths.

a) OPM simulation with sensor stand-off of 3 mm. Coloured lines represent mean model evidence difference values when *independent* sensor location error is added. Model evidence differences increase with increasing source strength (along x). Up to 5 mm or error (green line), the increased SNR afforded by decreased sensor distance enables detection sources which are 20 nAm or more in strength (never possible with SQUIDs, see subplot b). Black line marks the significance threshold of 3 where the combined model is significantly (~20 times) more likely than the cortical model. **b)** SQUID simulation with sensor stand-off of 30 mm. Coloured lines represent mean model evidence differences when independent location error is added. Larger distances to the simulated sources means that higher source strength is needed for model evidence differences to be significant (10 nAm in panel a versus 50 nAm in panel b). Due to the relatively low SNR, only small errors are tolerated by the SQUID system: errors above 2 mm rule out detection of hippocampal sources. **c)** OPM simulations, same format as panel a but dotted lines showing effects of *systematic* errors as source strength increases. Independent errors (solid lines) included for reference. Systematic errors generally give higher model evidence differences and so are less detrimental to inferences about hippocampal sources than independent errors (but this difference is not significant). Syst systematic, indep independent. **d)** SQUID simulations, same format as c. Average model evidence differences when systematic error is added are shown with dotted lines. Again we found the difference between systematic and independent errors across all noise levels to be positive but not significant.

Next we were interested in the relationship between increased sensitivity to hippocampal signals, and increased sensitivity to error. That is, do we need more accurate models to make use of OPM data, and how accurate must the modelling be in order to benefit from having the sensors closer? To first visualise an answer to this question, we replotted a subset of the data shown in Figure 5.3 to show the effects of location error on model comparison (**Figure 5.4a,b**). We chose 10 and 50 nAm because these were the approximate source strengths at which the OPMs and SQUIDs could detect hippocampal sources (without error). We found that the model evidence difference falls off faster as a function of error with OPMs than with SQUIDs (compare grey lines, **Figure 5.4a,b**), and that at approximately 10 mm of error there is no longer an advantage of having the sensors closer, as the difference between models is below the significance threshold of 3 regardless of source strength.

As the model comparison on its own may give false positive (or negative) results by chance if the underlying models do not differ in their (actual) ability to explain the data

parsimoniously, we calculated the Bayes Omnibus Risk (BOR) to ensure that this was not the case (**Figure 5.4c,d**). We did this for both source strengths and both error types. BOR was thereby used to quantify the reliability of the model evidence differences as a function of error. We found that although the BOR values mirror the model evidence difference values well generally, they suggest a slightly more conservative interpretation of the results. For example, with SQUIDs the model evidence differences at 0, 1 and 2 mm are (just) significant (**Figure 5.4b**) but the BOR values are between 0.08 and 0.09 (i.e. not significant). Per convention we define the BOR significance threshold to be 0.05. Thus, at low errors and 50 nAm sources detection of hippocampal sources is not reliable using SQUIDs.

Conversely, we find that with OPMs, the BOR stays below the significance threshold up to and including 5 mm of error for 50 nAm sources (grey lines, **Figure 5.4c**). However, at 10 mm of error, although the model evidence difference for the systematic error condition is significant (**Figure 5.4a**, grey dotted line), the BOR suggests that these two models may be equally likely.

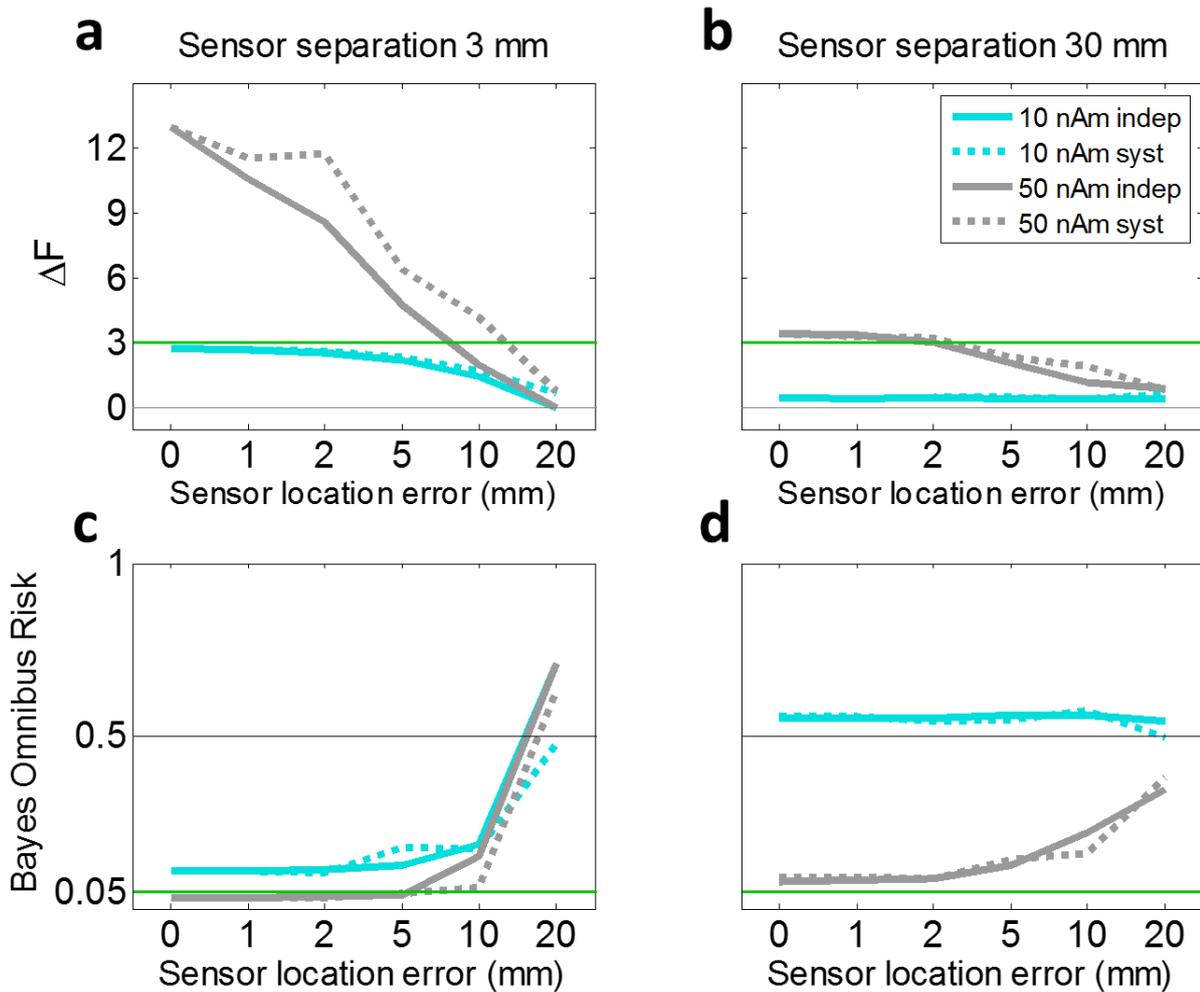


Figure 5.4 Evidence for source models as an effect of sensor location error

a) OPM simulation with sensor stand-off of 3 mm. As sensor location error increases (along x), model evidence differences rapidly decrease. This is most pronounced when source strength is higher (50 nAm; grey lines). High SNR entails high sensitivity to error. Green line marks significance threshold of 3 (combined model 20 times more likely than cortical). b) SQUID simulation with sensor stand-off of 30 mm. Evidence for the correct source model is lower (almost no model evidence difference at 10 nAm, blue lines), and decreases less rapidly. SQUIDs are less sensitive to activity and more robust to errors. c) Bayes Omnibus Risk (BOR) as a function of location error when sensor stand-off is 3 mm. With 50 nAm sources, model evidence differences are significant with up to and including 5 mm of error whereas 10 nAm sources never give a reliable (or significant, a) difference. Green line marks significance threshold 0.05 where risk that model frequencies are equal is 5%. d) BOR as a function of location error when sensor stand-off is 30 mm. Even when model evidence differences are significant at 0-2 mm with 50 nAm source (grey lines, b), the BOR probability that the two models are equally likely is not significant.

Effects of independent and systematic errors in sensor orientations

In order to characterize more rigorously what the empirical requirements for OPM-based detection of hippocampal signals might be, we next investigated what the effects of sensor orientation errors were. Again we tested the effects of both systematic and independent errors on model evidence differences when simulated sources were hippocampal. Similarly to with location errors, this analysis revealed that systematic and independent orientation errors negatively affect our ability to detect hippocampal sources, can be compensated for to an extent by increased source strengths, and are largely equally detrimental to model comparison (**Figure 5.5**). The differences between independent and systematic errors were less pronounced than with location errors and also not significant with neither OPMs (BOR=0.604), nor SQUIDs (BOR=0.653). We also found a similar difference in tolerance of errors as before: OPMs give significant model evidence differences in the face of larger errors and lower source strengths than SQUIDs. While both 5 and 15° errors can be tolerated by OPMs when sources are <20 and <50 nAm respectively, SQUIDs require 100 nAm source strength to detect hippocampal sources when there is 5° of error, and can never detect hippocampal sources when there is 15° of error, regardless of how strong the source is (**Figure 5.5**).

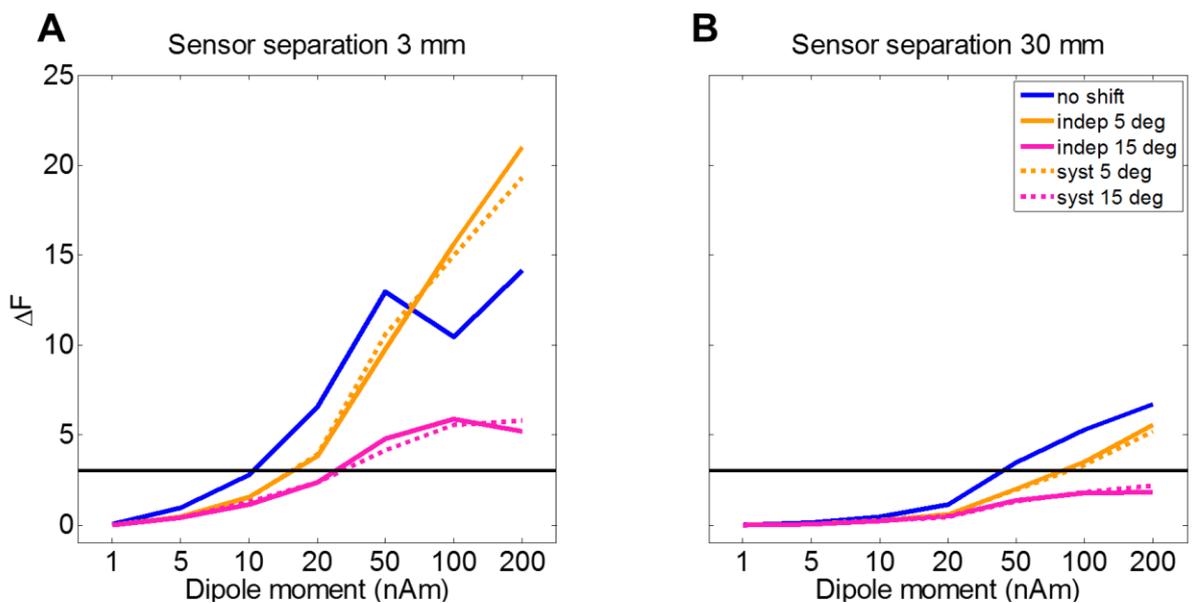


Figure 5.5: Effects of systematic and independent sensor orientation errors on ΔF across simulation strengths.

Same format as Figure 5.3 but with perturbed sensor orientations rather than locations. Independent perturbations are approximately equally detrimental as systematic both when relatively large (15°) and relatively small (5°). **A)** OPM simulation with sensor stand-off of 3 mm. At 5° of sensor orientation error (orange lines), it becomes possible to distinguish between models at <20 nAm source strength, whereas >20 nAm is required when errors are 15° . **B)** SQUID simulation with sensor stand-off of 30 mm. Sources must be 100 nAm to be detectable when there is 5° of orientation error added to the sensors. Conversely, is not possible to reliably infer hippocampal activation when there is 15° of orientation error, even when the source strength is high (200 nAm, compare to ~ 20 nAm required with an OPM array, A). Syst systematic, indep independent.

We then investigated the relationship between sensitivity to hippocampal signals versus modelling error in the context of sensor orientation uncertainty. Again, we asked, how accurate must the modelling be (in terms of sensor orientation), in order to benefit from having the sensors closer? Similarly to **Figure 5.4**, we first replotted data shown in the previous figure to assess the effect of error on model comparison with 10 and 50 nAm sources (**Figure 5.6a,b**). Here, we observe a similar pattern as before, namely that the model evidence difference falls off faster as a function of error with OPMs than SQUIDs. Interestingly, we also find that although 50 nAm sources on average give highly significant model evidence differences with OPMs when the orientation error is 5° (10.58 and 9.75 for systematic and independent errors respectively, **Figure 5.6a**), the associated BOR values are not significant (both 0.24). The same relationship is found with 15° errors: model evidence differences are significant at 4.79 and 4.16 for independent and systematic orientation errors, while BOR values are likewise not significant at 0.25 and 0.26 respectively (**Figure 5.6c**). This means that even with very small (5°) errors, the high SNR obtained by using OPMs does not afford hippocampal detectability. We also found that with orientation errors added to SQUID sensors (included for comparison, but practically does not occur when sensors are arranged in a fixed configuration inside a scanner **Figure 5.6b,d**), it is never possible to detect hippocampal sources as simulated here.

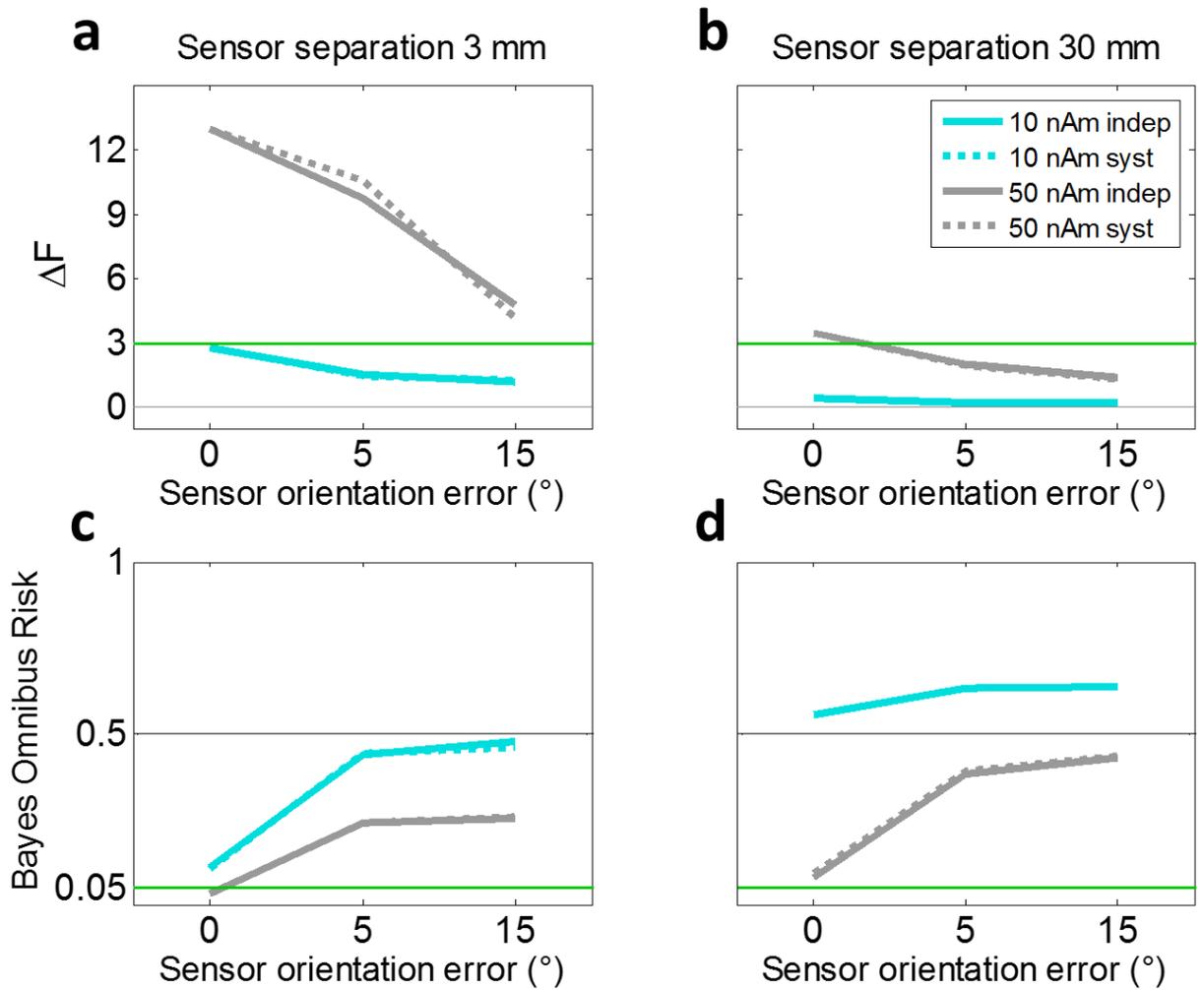


Figure 5.6 Evidence for source models as an effect of sensor orientation error

Same format as Figure 5.4 but showing model evidence differences and BOR values as a function of sensor orientation as opposed to location errors. **a**) OPM simulation with sensor stand-off of 3 mm. As sensor orientation error increases (along x), model evidence differences rapidly decrease. Again this is most pronounced when source strength is higher (50 nAm; grey lines). Green line marks model comparison significance threshold of 3 (combined model 20 times more likely than cortical). **b**) SQUID simulation with sensor stand-off of 30 mm. Evidence for the correct source model is lower with either source strength - there is almost no model evidence difference with 10 nAm sources (blue lines). SQUIDs are less sensitive to activity. **c**) Bayes Omnibus Risk (BOR) as a function of location error when sensor stand-off is 3 mm. With 50 nAm sources, adding either 5 or 15° of orientation error invalidates the model evidence difference. Green line marks significance threshold 0.05 where probability that model frequencies are equal is 5%. **d**) BOR as a function of orientation error when sensor stand-off is 30 mm. In all conditions shown here, the BOR is not significant.

Effects of sensor error on the combined generative model alone

So far we have looked at model evidence differences between the combined and cortical models. However, a potentially more sensitive measure of how modelling errors affect our ability to detect hippocampal sources might be to assess changes in model evidence values for the combined model alone, as a function of error. We reason that the optimal condition for measuring hippocampal activity is with zero error, and therefore assess how error affects the model evidence we obtain with the combined model when modelling error is added (**Figure 5.7**). This analysis revealed stronger effects of error than when the two different generative models were compared. For example, at 1 mm of systematic location error, the model evidence value for the combined model is significantly decreased in both the OPM and SQUID simulations, specifically by 26.7 and 10.61 log units respectively (**Figure 5.7a,b**). This equates to the combined model being more than 390 million and 40,000 times more likely for the zero-error condition.

Several other observations emerged from plotting the results in this way. First, we found that generally, model evidence values are more strongly affected by modelling errors when the sensor array is closer to the head (compare left and right columns, **Figure 5.7**). Second, we found a tendency for systematic errors to be equally or less detrimental in the context of location error, but more or equally detrimental in the context of orientation error (compare upper and lower panels, **Figure 5.7**).

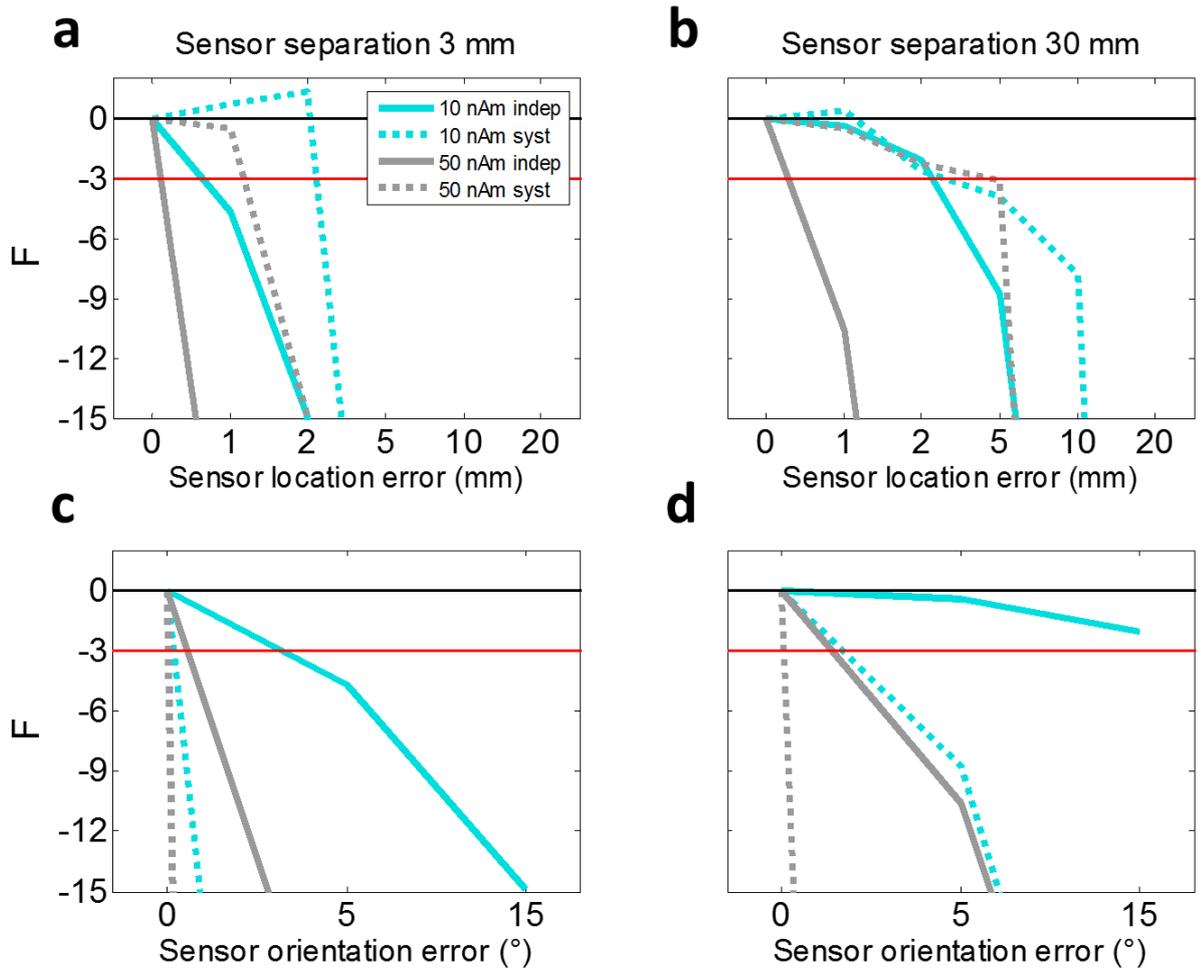


Figure 5.7 Evidence for combined generative model as sensors degrade

Change in model evidence values of the combined generative model as a function of sensor errors. **a)** Model evidence for the generative model in an OPM array as a function of sensor location error. Model evidence values normalised to zero error condition (marked by black line). Red line marks significance threshold of -3 where the zero error model is 20 times more likely than the one in comparison to it. Systematic errors (dotted lines) are less detrimental to the model evidence: values drop off less sharply with error and stay above the significance threshold until 1 and 2 mm for 50 and 10 nAm sources respectively. The stronger (50 nAm) source causes the model evidence to drop off more sharply with error than the weaker (10 nAm) source in the independent location error condition where even 1 mm of error causes a significant decrease in model evidence at both source strengths. **b)** Same format as a, but for a SQUID sensor array. The effect of sources being stronger is the same as sensors being closer to the head (panel a): more rapid drop-off of model evidence with error. Both error types tolerate up to 2 mm of error when sources are weak (10 nAm, blue lines) or error is systematic but source is strong (50 nAm, grey dotted line). **c)** Model evidence values for the combined generative model as a function of orientation errors when sensors are 3 mm from the scalp. All

model evidence values decrease significantly from only 5° of error, regardless of type or source strength. Systematic errors cause larger decreases than independent errors. **d)** Model evidence values for the combined generative model as a function of orientation errors when sensors are 30 mm from the scalp. While model evidence for the combined model does not significantly decrease when sources are 10 nAm and errors are independent and up to 15°, all other conditions tested show a significant decrease when only 5° of error is added. Again, systematic errors are generally more detrimental, although the 10 nAm trajectory is approximately equal to the 50 nAm independent error trajectory.

Discussion

In this section, we introduced a simple simulation framework to evaluate the advantages and limitations of detecting hippocampal sources using OPMs placed directly on the scalp. It appears that OPMs provide the potential for approximately 5-fold sensitivity improvements if testing specific hypotheses about deeper structures such as the hippocampus (**Figure 5.1**). This is higher than that reported when detectability is determined on the basis of beamformer reconstructed time courses (1-2 fold) (Boto et al., 2016). Thus, we posit that a model comparison approach may enable further increased sensitivity through more precise priors.

However, increased sensitivity to sources implies increased sensitivity to modelling error. We found that while small modelling errors have large negative effects on OPM inversions, this is to an extent offset by the increased SNR afforded by having the sensors closer to the scalp (**Figures 5.3-5.6**). Importantly, because OPMs can be placed flexibly, they are more prone to both sensor orientation and location errors. These may occur either systematically or independently (or both). We found that all errors are highly detrimental to hippocampal detectability, confirming our hypothesis that increased sensitivity also extends to errors. This supports the notion that model-comparison based methods may be more sensitive to deeper sources, as the effects of modelling errors reported in other simulations (Boto et al., 2016) show that the benefits of OPMs for deep sources are negated by forward modelling errors of ~20% (discrepancies between simulated and inverted lead fields).

Our results suggest that OPM sensors must be physically constrained in order to both make use of the stronger signal, and maximise their potential. In particular, modelling the orientations of the sensors (which is not an issue with SQUIDs where the orientation cannot vary) is very detrimental to successful model comparison-based detection. We find that the empirical error should be kept as low as possible, at least

$<5^\circ$ (**Figure 5.6-5.7**). We therefore propose that the sensors are oriented and fixed with respect to the head, eliminating both independent and systematic errors *a priori*. This could be done with a head-cast like design (Chapter 3) but where the head-cast itself is (more) rigid and contains slots which fix both the orientation and location of the sensors with respect to the head, and each other.

It is worth noting here that although unclear from these simulations what the exact source-level SNR improvements may be, and whether they could be improved further by arranging sensors in a more data- or anatomy-driven manner. It is also unclear what the optimal sensor coverage density might be, although simulations suggest that the improvement continues to increase with density (Boto et al., 2016).

Finally, although it might be informative for designing an acquisition device, this simulation set-up is inherently optimistic and potentially unrealistic for several reasons: sensors are located in the throat and face and SQUIDs are located outside a dewar system. Here we simulate OPMs as magnetometers but axial gradiometers have recently been configured and empirically validated (but are not yet fully implemented in the commercially available devices). Nonetheless, we posit that these findings are useful as demonstrations of the sensitivity and flexibility of these new sensors, as well as novel guiding principles for optimisation of acquisition protocols for these new devices.

Key points

- Reducing the distance between the brain and sensors gives ~5 fold sensitivity improvement for hippocampal sources.
- Increased sensitivity is accompanied by increased sensitivity to errors which may be offset to an extent by increased SNR.
- A method for locating and orienting the sensors with respect to the head is needed in order to eliminating co-registration errors while minimizing the scalp-sensor distance.

This chapter does not derive from a manuscript in preparation.

Chapter 6

Discussion

General Discussion

Overview

The aim of this thesis was to advance acquisition and analysis methods for detecting signals from the human hippocampus using MEG. In doing so, an extension of the standard generative model used for inversion of MEG data was created by adding a hippocampal mesh. This extended model was then tested across a range of simulated conditions (Experiment 1). Next, the development of a new prototype of flexible head-casts used for stabilizing and accurately repositioning the head in the MEG scanner was described and tested empirically (Experiment 2). The head-casts and hippocampal source model were then successfully applied to acquire and analyse real data and revealed hippocampus-specific activity (Experiment 3). Moreover, the detectability of hippocampal sources with a novel sensor type was characterized in simulations (Experiment 4).

In this chapter, this work is considered in relation to other recent work in the relevant field(s), and a new perspective on the ongoing debate about whether it is possible to detect hippocampal activity using MEG will be presented. Then, a set of new potential research questions made feasible through the combination of the methods developed here are suggested, along with and in relation to potential analyses and clinical applications. Finally, a 3D printing based head-cast design for OPM sensors is presented.

Chapter 2: What is needed to see hippocampal activity in MEG?

In Chapter 2 we considered how to test for hippocampal contributions to the MEG signal using simulated data with an anatomical and electrophysiological source model of the hippocampus. The underlying rationale was that if a source is hippocampal, then inclusion of a hippocampal mesh in the generative model will give a more parsimonious inverse solution. Therefore, by comparing a generative model with a hippocampal mesh to one without, it should be possible to infer the presence of a hippocampal source based on the difference in model evidence obtained with different models applied to the same data. Thus, we used a Bayesian framework to make probabilistic statements about (simulated) hippocampal activity through model

comparison. We found that it is possible to detect hippocampal sources when the co-registration error is low (<3 mm) and SNR modest (>-20 dB).

We also tested the specificity and sensitivity of this model comparison. We found that the specificity is high (or false positive rate is low), as the model evidence difference between the models is not significant when sources are cortical. We also found that the sensitivity is high (or false negative rate is low) as it is possible to detect hippocampal sources based on the model evidence differences (**Figures 2.5 and 2.8**). The sensitivity depends on the simulation parameters however and decreases with modelling errors.

An important assumption made in this study was that a representative model of hippocampal sources could be constructed based on the pyramidal cells distributed in macro-columns of Ammon's horn (hippocampal subfields CA1-4). We used the external tessellation of the hippocampus as to model and simulate the sources. It should be noted however that the hippocampus also includes the subiculum and dentate gyrus subfields, the latter of which interlocks with Ammon's horn. Nonetheless, pyramidal cells which are thought to constitute the main component of the measured neocortical MEG signal are mainly found in Ammon's horn. Our model is therefore a first approximation but it would be useful to consider higher resolution MRI images obtained with 7T for segmentation of the CA1-4 and dentate gyrus subfields (Wisse et al., 2012).

Another recent study which also assessed source localisation accuracy of hippocampal sources in simulated data used somewhat comparable methods (Attal and Schwartz, 2013). For one, a similar electrophysiological model of the hippocampus was used. For another, the authors simulated simultaneous activations as well as variable temporal overlaps of hippocampal and cortical activations. They found that while it was not possible to detect hippocampal sources when cortical and hippocampal sources were simultaneously active, it was possible to do so when a neocortical source was no more than 50% of the hippocampal in amplitude. In these cases, hippocampal sources were detected but a local maximum was created in the thalamus with dSPM and sLORETA algorithms (Attal and Schwartz, 2013). We also simulated hippocampal sources twice the strength of cortical sources. In our results however, we found that the model evidence difference was significant in 23% of cases where there were two hippocampal and two (temporally overlapping) cortical sources. This result, however, compares favourably well to another study, where Stephen and colleagues showed in simulations that MEG is able to correctly localise activity from

the hippocampus and parahippocampal gyrus only when there is no temporal overlap (Stephen et al., 2005). However, the ability to reliably detect only 23% of hippocampal sources in the presence of concurrent activity is a non-trivial problem which, if true with empirical data as well, suggests that the modelling is limited in its usefulness.

Moreover, in the simulation study by Attal and colleagues, the hippocampal tail and edges were found to be more prone to localisation errors (Attal and Schwartz, 2013). We did not test for this explicitly but it would be interesting and possible to do so both with Free energy and/or cross validation error calculations. One possibility for doing so would be to change the source model. For example, one might use the rotation method presented in Chapter 4, take out different portions of the hippocampal mesh (e.g. the most anterior 25%), swap the hippocampi of different subjects, change their sizes, etc. It would in itself be informative to compare the performance of these altered models both for simulated and real data.

Like other simulation studies looking at hippocampal activity, we also simulated dipoles with some spatial spread (approximately 1 cm). Others have used patches of varying size (Attal and Schwartz, 2013; Chupin et al., 2002). The use of patches of varying size has mainly demonstrated a detection saturation effect with size of the patch, due to the curvature of the hippocampus which causes sources to oppose each other and cancel out. We would most likely observe the same if we were to simulate more extended patches.

Another potentially fruitful extension to this framework would be to use a different method for computing the forward head model. We used a single shell but would likely benefit from having a more realistic model with multiple compartments such as a Boundary Element Method-based one (Brebbia and Dominguez, 1989; Stenroos, 2016).

We note here that while we evaluated our two anatomical models in a Bayesian context, the models and the framework are not inherently limited to this approach. For example, we also assessed the dipole localisation error obtained with the different models and found this to echo the Free energy findings. It would be interesting to examine what the model comparison would look like in other model comparison regimes or with different inversion parameters. There are both similar metrics to Free energy such as the Bayesian Information Criterion and Akaike's Information Criterion which could be evaluated (as evaluated in Penny, 2012), as well as more conventional methods which could take direct advantage of the known ground truth, such as

localisation error, point spread, and cross talk (as evaluated in Attal and Schwartz, 2013).

Overall, the usefulness of the approach and the validity of simulation results always depend on empirical validation as well as on additional assessments of any possible impacts from relevant (in this case for example physically nearby) potential influences. Hence, one potential caveat of this study or model is that it might be helpful to model more of the surrounding structures in order to directly distinguish between contributions, and quantify this probabilistically. For example, the amygdala as well as the entorhinal and retrosplenial cortices are strongly coupled to the hippocampus functionally and anatomically. For the basolateral nucleus of the amygdala and the basal ganglia and related structures, the lack of laminar and oriented cell structure suggests that distributing the dipoles inside regular volume grids rather than on the surface would be more likely to capture true sources. It follows that more simulation and empirical work could be done on assessing the separability of these sources. However, in our multiple dipole approach where hippocampal dipoles were simulated to be twice as strong as neocortical ones (which in turn raises the question of what an appropriate current density would be for other deep structures), we only found modest model evidence improvements.

Chapter 3: Where is the brain?

In Chapter 3 we focused on developing a method for improving co-registration and head stabilization. The flexible head-casts developed are an extension of a previous prototype (Troebinger et al., 2014b). The aim of the present PhD work was to minimize uncertainty in the forward model, while maximising subject safety and comfort, including avoiding obstructing vision. Enabling vision greatly increases not only the comfort of the subjects, but notably also the usefulness of these devices, as cognitive, psychophysical, as well as purely visual experiments are become possible. Subjects are more comfortable and less likely to experience anxiety when vision is not obstructed. We also made the construction more practical and accurate by using an anatomical MRI image as the basis of the cast, as opposed to using an optical scan. Critically, the performance of the new head-casts was consistent with, or better than the previous prototype and both safety and comfort were improved: we found the within- and between-session movement levels to be $<.25$ and <1 mm respectively. An important feature of the head-casts generally is that they do not rely on self-stabilisation by the subject which avoids muscle artefacts which are likely to

contaminate the signal (Adjamian et al., 2004; Kumar et al., 2003; Muthukumaraswamy, 2013; O'Donnell et al., 1974; Whitham et al., 2007).

We tested and quantified the stability and reproducibility of the head position within and across sessions, and demonstrated high spatial reproducibility of source-level results. However, we did not carry out a direct comparison with non-head-cast data here. This has since been done by an independent group and shown to give increased stability of beamformer estimated source orientations (Liuzzi et al., 2016). In our case, such a direct comparison might have enabled us to directly quantify the improvements when all other acquisition parameters were kept constant in the context of the measurements presented here.

Nevertheless, the new flexible head-casts provide the potential for increasing the SNR through repeating the same experiment many times and no longer being limited by head movement which in turn opens up new and exciting possibilities. For example, would it be possible to formulate hypotheses based on more subtle signal features than before (such as thalamic, lateral geniculate nucleus, or brainstem signals)? Would it be sensible to build up very high SNR datasets using only a small groups of subjects, and conduct analyses at the single subject - rather than at the group - level? In doing so, the focus would shift from group-level findings to single subject features.

Thus, the next question is whether these head-casts should provoke new ways of thinking about data analysis and pipelines. For example, an interesting and counterintuitive finding is that beamformers perform more optimally when there is some noise in the data, if the source model is not completely accurate (Hillebrand and Barnes, 2003). Therefore, if there is less noise and a more accurate model, should some error be added back if using beamformers? Or, one might ask whether other algorithms such as multiple sparse priors (MSP) would be more suitable for high SNR data (Friston et al., 2008a)? Research on the opposite question has been carried out, essentially asking how best to account for the uncertainty in probabilistic terms (López et al., 2014, 2012). Similarly, it has previously been shown that it is not worth the extra effort of building a subject-specific forward model for inverting MEG data under normal circumstances; that a canonical mesh is just as good given the expected modelling uncertainty inherent in MEG data (Henson et al., 2009). However, it has also been shown – also using Free energy – that when co-registration errors fall below 5 mm, it is worth creating a subject-specific forward model, as this gives higher model evidence than a canonical model does (Troebinger et al., 2014b). Thus, it is worth constructing subject-specific forward models for head-cast data analysis, and it is worth considering

- and ideally formally testing - which inversion algorithms are most appropriate for achieving the most accurate interpretation of the data.

This leads to the next point, which is how to find and make use of the spatial and structural information available in the MEG data? One major research field concerns itself with distinctions between the different laminar layers of cortex, the interactions between them, and the roles these interactions play in cognition. Until now, it has only been possible to explore and study these layers in animal models, fMRI, or invasively. However, if head-casts make it possible to carry out non-invasive laminar electrophysiological studies in humans in vivo, with high temporal resolution, then it would be a key method for bridging and expanding these neuroscience findings in computational terms. There are already promising MEG studies which address this issue (Michalareas et al., 2016; van Pelt et al., 2012). The main distinction of interest is formulated in terms of a predictive coding account of human cognition. Specifically, bottom-up (or feed-forward) processes are thought to originate from pyramidal cells in superficial layers (layers II/III), whereas top-down (or feedback) processes are thought to originate from pyramidal cells in deeper layers (layers V/IV). Due to the frequency differences between the two directions, MEG lends itself well to human investigations. While bottom-up processes have been shown to occur at higher frequencies such as in the gamma band (30-90Hz), top-down processes occur at lower frequencies such as in the beta band (15-30Hz) (Bastos et al., 2012).

Another potentially fruitful application is within decoding analysis, another relatively new area of MEG research (Cichy and Pantazis, 2015; Jafarpour et al., 2013; Myers et al., 2015; Stokes et al., 2015). The basic idea is that if it is possible to decode spatially overlapping features of the signal by distinguishing between different spectral properties of the signal (e.g. activity in different frequency bands). One would expect that data with less head movement and higher SNR would be ideally suited for decoding analyses, as the rich spectral content is more consistent over time, and the dipoles stay oriented in the same way if the head does not move (Liuzzi et al., 2016).

Perhaps the most obvious use of the new head-casts is for questions related to changes over longer time than during a single scan. Such use spans from long term longitudinal studies which could be used to study changes in brain structure, function, and spectral signatures over years or decades, to memory experiments depending on consolidation of newly acquired knowledge over short timescales (e.g. one hour, one night's sleep) to longer timescales (e.g. days, weeks). While the dynamics of encoding and retrieval of information have been studied using MEG (Crespo-García et al., 2016;

Grimault et al., 2014; Kaplan et al., 2012a), they have not been studied in the context of longer times than single experiments. Generally, longitudinal experiments would be well-suited for using the head-casts, both because SNR can be built up over time with a large number of trials, and because accurate re-positioning of the head inside the scanner is key to doing so, and to making valid inferences across recording sessions.

Apart from memory experiments, three other particularly well-suited areas of research for longitudinal MEG experiments are developmental, degenerative, and pharmacological neuroscience. Child development is rarely studied in MEG mainly because children often move their heads, which is moreover made easier by the fact that they have heads which are far smaller than those of adult subjects which the current MEG dewars are designed to hold. Thus, stabilizing the head opens up an exciting possibility for studying the temporal dynamics of child and adolescent development with high spatial resolution (unlike EEG). Naturally, with growth would also come the need for updated head-casts and anatomical brain images for accurate modelling. With degenerative disorders, especially those which affect movement control, it has previously been difficult, if not impossible, to detect changes using MEG. Although head-casts potentially open up this possibility and thereby hold the promise of potentially facilitating identification of clinically useful trait- and state-markers, it is important to uphold and extend the safety regulations we propose. Not only would it be more challenging from the experimenter's perspective to scan elderly and/or sick patients, it would also potentially be unsafe, as well as challenging for the participant to get in and out of the dewar unaided. Therefore, although head-casts hold potential for characterizing and disambiguating brain changes in relation to degenerative disorders, translating the technology into a useful tool for these populations carries with it significant challenges. Finally, it would be interesting and potentially more straightforward to carry out more pharmacological studies in MEG and make use of the increased SNR. For example, drugs with longer half-lives could be administered and the participant could be scanned at regular intervals (e.g. for 15 minutes every half hour) to track slower changes and their effect on resting state, induced, and/or evoked responses. This would give way for a deeper understanding of the relationships between behavioural and neural changes in responses to certain drugs and combinations hereof.

Chapter 4: Is it really the hippocampus?

The results from Chapter 4 demonstrated hippocampus-specific activity recorded with MEG. The data was acquired by asking subjects to perform a cognitive task known to evoke hippocampal theta band oscillations (Bush et al., 2015) while they were wearing a head-cast. Through application of Bayesian model comparison and cross-validation, we found that lateral rotations of the hippocampal portion of the generative model significantly decreased its predictive power, even when these errors were as low as 5°.

As with the simulation study, this leads to a question of hippocampal model specificity. Would the results hold if a given hippocampal mesh originated from another subject's brain? In other words, how sensitive are we to the individual subject's anatomy and would it be possible to distinguish between anatomical models on the basis of the MEG data? Alternatively, one might ask whether a canonical hippocampal mesh can capture and explain hippocampal sources? One might imagine a canonical hippocampal mesh similar to the canonical cortical mesh implemented in the statistical parametric mapping (SPM) software where the size and shape are warped into an individual subject's forward model so as to fit most appropriately and be used as a set of spatial priors. One reason to expect that a canonical mesh might be sufficient is that there is less structural variability in deeper than more superficial brain structures. It would be relatively straightforward to test both the hypothesis that the MEG data can be used to distinguish between anatomical models, and/or that a canonical model is good enough using model evidence (Henson et al., 2009; Troebinger et al., 2014b). In this case, one would also obtain a direct quantification of the improvement from using a subject-specific model (if any), perhaps as another form of measuring the amount of spatial information in the MEG data. Further, regardless of whether the hippocampal model used for analysis was canonical or not, it would also be relatively straightforward to configure an automated way of testing the sensitivity to rotations of the hippocampal model, as was shown in Chapter 4. This might then constitute one method of assessing hippocampal engagement. Other questions which might be asked using the nested hippocampal mesh and structural manipulations or omissions hereof are: does the activity lateralise to one hippocampus (previous literature suggests that right-lateralisation might be expected (Iglói et al., 2010; Jacobs et al., 2010)? If so, is this lateralisation consistent across subjects? This would be testable by omitting one hippocampus in one generative model and the other in another, and comparing model evidence values. Other questions include: is there evidence for an anterior-posterior dissociation? Do certain features of the signal localise to certain

regions or subfields of the hippocampus (e.g. frequency, amplitude, or phase of an induced response)?

One implication of the results presented here is that the simplifying assumptions made in the simulation paper are valid – if they were not, the empirical replication would not work. The most important of these simplifications, which turned out to produce valid results, is the hippocampus pyramidal cell modelling in the form of a tessellation of the surface envelope.

A prediction made from the simulation results and an obvious question to ask in relation to the data collected is whether the use of head-casts is truly necessary or whether the modelling approach would reveal the same levels of sensitivity with non-head-cast data? As the premise here is combining the methods presented in previous chapters, this would be relevant to formally test. This is an empirical question as the spatial resolution of (normal) MEG is relatively high (~5-8 mm) (Brookes et al., 2010), depending on how and how much data is collected. If head-casts are not necessary for detection of hippocampal sources (or, the sensitivity was only slightly lower, e.g. lateral rotations of 15° were significantly worse than the true model), then the model-based analysis might have further reaching relevance and could be easily and immediately implemented in a far wider context.

Regardless of dependence on head-casts, we believe that the empirical, experimental demonstration of hippocampal detectability achieved provides a novel method for formally assessing directly the contribution of hippocampus to non-invasively measured signals. Most excitingly, this result in turn implies that we can begin to bridge electrophysiological findings from rodent hippocampal studies with human neuroscience, ask more complex questions through more complex tasks, and thereby probe new nuances of behaviour, memory, experience and learning. Understanding spatial cognition in particular sets out a basis for better understanding many other cognitive processes and mechanisms such as decision making and social relationships (Barron et al., 2013; Eichenbaum, 2015; Maguire et al., 1999; O'Keefe and Nadel, 1978). Such an improved understanding would also allow identification of features and functions in humans which would aid determination of the homologies between human and rodent hippocampi at a systems level. New, specific and testable hypotheses can be formulated for MEG to address many of these open questions.

As a starting point, in addition to the methodological modelling related questions listed above, other, more cognitive questions could be asked of the data collected with the present experiment. For example, one would expect that novelty of the stimulus should

have an effect on theta: based on previous literature, one might expect increased theta power during retrieval (object presentation) for objects that are novel versus those which are familiar (Kaplan et al., 2014). Specifically, we predict that this effect will be most prominent in anterior portions of the hippocampal body (Kaplan et al., 2014; Viard et al., 2011). Note that “novelty” is used as a relative term here: as each set of objects is presented twice (in two different spatial environments), the data could be split into first retrieval trial (“novel”) and contrasted with the second retrieval trial. Another feature we could potentially measure is a neural correlate corresponding to the strength of memory encoding. Previously, theta power has been shown to correlate with subsequent memory performance both in non-spatial (Osipova et al., 2006; Sederberg et al., 2003; Serruya et al., 2014), and spatial (Kaplan et al., 2012b) experiments. This effect is likely to be found mainly in the posterior portion of the hippocampus (Doeller et al., 2008; Kaplan et al., 2012b; Nadel et al., 2013). It would be particularly interesting to investigate this question, as new data suggests that the opposite is true; that the theta power decreases during item-place encoding (Crespo-García et al., 2016). Factors such as the exact frequency band chosen (Jacobs, 2014) as well as task parameters are likely to play an important (but thus far unclear) role. It is therefore useful to consider neither the theta or gamma band activity as unique and separate markers of brain activity, but rather to consider each as a single data feature in single dimensions of analysis. Interactions with other frequencies and brain regions for example may be equally if not more informative about the underlying processes, e.g. phase coupling with medial prefrontal cortex as shown in a very similar version of this task (Kaplan et al., 2014). In addition to cognitive questions such as these, interesting computational questions could likewise be raised. For example, predictions made from computational models of spatial memory (Byrne et al., 2007; Byrne and Becker, 2004) might include specific hypotheses related to retrosplenial cortex’s role in translation between egocentric and allocentric reference frames, transmitting this information to the hippocampus through a dorsal pathway which combines this input in the posterior parahippocampal gyrus.

Chapter 5: The coolest magnetometer is not the best one

In Chapter 5 we examined a set of empirical factors in using new optically pumped magnetometers (OPMs) for MEG, specifically with the aim of measuring hippocampal sources. We were interested in assessing how to most efficiently use these new sensors. We found that as with standard MEG, co-registration is detrimental to our

ability to detect hippocampal sources. Critically with OPMs, this can take a new form as the sensors are not fixed with respect to the others so the sensors can now move with respect to each other and vary in both location and orientation. We also found that while OPMs are ~5 times more sensitive to hippocampal signals because they are closer to the source, they are also more sensitive to errors (**Figure 5.5**). An interesting extension of these simulations would be to assess the effects of independent and systematic sensor orientation or location errors in combination with each other. Here, we only assessed these effects in isolation but they could likely co-occur in reality and potentially interact in non-linear ways. Due to the results and this prediction, we therefore propose using a modified version of the head-cast design to orient and stabilize the sensors with respect to the head. This removes co-registration to the best of our ability, as the relationship between the sensors and brain anatomy is known a priori (without needing localisation of fiducial coils).

Although simulations assessing OPMs have not been carried out in great numbers (yet), a highly relevant paper for comparison with the results presented here has recently been published (Boto et al., 2016). The authors found 5-10 fold improvements in signal amplitude for superficial sources and lower, approximately 2 fold improvements for deeper sources. This was quantified as ratio of Frobenius norms of the forward field vectors for the different sensor types. This result raises an important albeit subtle point: moving sensors closer to the scalp will produce a shift in the relative SNR of the hippocampal and cortical. This means that although the SNR of the hippocampal sources increases as a consequence, the ratio of cortical to hippocampal SNR also increases meaning that there is a risk that hippocampal detectability could decrease overall. The authors of the study did not explicitly test hippocampal detectability but it would be interesting to expand the simulations presented here to include cortical sources, and examine the relative contributions to the signal and begin to explore how best to account for this. Another interesting analysis was direct quantification of the source-level SNR for comparison between SQUID and OPM set-ups. This direct SNR quantification avoids the issues related to comparing different sensor types (i.e. magnetometers versus gradiometers). Nonetheless, the SNR calculations also allow comparison of arrays which are different in other ways, e.g. sensor number/density or location/distribution. The authors also assess the effect of error on OPMs versus SQUIDs and similarly to here, find increased sensitivity with increased SNR. Specifically, they show that a 5% error in the forward field used for inversion (relative to the one used for simulation) eliminates the advantage of having OPMs when dipoles are shallow (<4 cm from surface). With

deeper dipoles (4-6 cm from surface), ~20% error is tolerated before OPMs reach the same performance as SQUIDs (Boto et al., 2016).

The simulations presented in this thesis await empirical validation, ideally using a subject also scanned with a flexible head-cast and tested on the spatial memory experiment (Chapter 4). The number of sensors required for reconstruction of hippocampal activity is an empirical (or simulation) question. A recent empirical OPM experiment using a solid head-cast showed that it was possible to reconstruct the source of the N20 median nerve response using a single OPM sensor placed in 13 different positions to simulate an array (Boto* et al., 2016). This is a promising and encouraging finding, as it also showed that for both evoked and induced activity, the sensor array is not limited to the number of sensors at hand. Due to the flexibility of the sensors and geometry hereof, one might also imagine more creative sensor placement for measuring hippocampal signals. For example, given the depth of the hippocampus, it is likely that portions of the magnetic fields generated would be best captured at locations outside of the standard MEG array, be it on the cheeks, neck, under the eyes, or inside the mouth or ears. Such arrays become possible with non-cryogenic sensors.

We hope that the simulation results and considerations presented here may be of use in guiding the design of OPM instrumentation and encourage the use of explicit hippocampal modelling in combination with the proposed head-cast design or variations hereof. This is catalysed by the significant decrease in both acquisition and maintenance cost, the possibility of placing the sensors in a subject-specific array with a higher sensor density, and the potential freedom from requiring a magnetically shielded room (if cancellation coils are developed to fully compensate for the Earth's field (Boto* et al., 2016)). The implications of this for cognitive and basic neuroscience have been discussed above and the clinical implications will be discussed below (*Clinical applications*). Further applications include several new domains of neuroscience including child and adolescent development (not previously well-studied in MEG because of small head size and large amount of head movement), social interaction whereby several subjects could be recorded simultaneously while wearing OPMs, movement and movement disorders or rehabilitation, spatial navigation while moving in the real world or in conjunction with virtual reality goggles. Many more unexplored domains likely exist.

Perspectives and Outlook

Clinical applications

In general, the hippocampus plays important roles in healthy cognition and is affected by a wide range of neurological, developmental, and degenerative pathologies such as Alzheimer's disease, Parkinson's disease, temporal lobe epilepsy, anxiety, depression, post-traumatic stress disorder, schizophrenia, and dementia (Baumgartner et al., 2000; Bisby et al., 2010; Brambilla et al., 2013; Eichenbaum, 2015; Gordon et al., 2013; Protzner et al., 2010; Squire et al., 2004). Additionally, it is hypersensitive to hypoxia and thus predisposed to impairments from lack of oxygenation, especially the CA1 region (Kreisman et al., 2000).

While other neuroimaging modalities such as fMRI and PET have demonstrated clear links between cognitive functions related to these disorders (such as memory) and hippocampal activity, the spatiotemporal dynamics of the hippocampal associated can only be revealed using a technique with higher temporal resolution such as MEG or EEG. The spatial resolution of MEG is better than that of EEG, making it more suitable for measuring hippocampal-specific signals. This in turn uniquely allows direct neuroimaging of the nature of hippocampal cognitive functions both in healthy and disordered states.

Perhaps the most direct translation of hippocampal MEG into a clinical context is in temporal lobe epilepsy. Epilepsy affects approximately 50 million people on a global basis, around 30% of which are resistant to drug treatments and may require surgical intervention (Zhang et al., 2014). Temporal lobe epilepsy accounts for about 60% of all epilepsy cases (Stefan et al., 2003) and is currently difficult to non-invasively image reliably pre-operatively. MEG is an ideally suited modality because epileptic events may in some cases last only few tens of milliseconds and rapidly spread from the seizure onset zone into complex networks. A promising new avenue is to assess and attempt to localize high frequency oscillations, particularly from the hippocampus and entorhinal cortex (Bragin et al., 1999). At present however, MEG is thought to be too costly and imprecise of an alternative to "presurgical surgery" where a rod or grid is implanted into the brain in the region(s) thought to contain the SOZ (Papanicolaou et al., 2005; Wennberg et al., 2011). Another obstacle is the need for the head to be still throughout recording which is not only difficult and tiresome to do over the course of many hours which is required for ictal activity to be observed due to the low frequency

with which it occurs. For these reasons, the preferred modality for telemetry (longer term observation) is EEG which can be used while the patient is moving around.

However, several facts suggest that OPMs hold great potential for substitution of intracranial electrodes. For one, they are far cheaper and therefore more accessible than a standard MEG scanner (Boto* et al., 2016; Shah and Wakai, 2013). For another, they will likely soon be compatible with telemetry and thereby improve the spatial resolution of the data recorded, potentially pushing towards good enough source localization to avoid the dangers, discomfort, and costs associated with pre-surgical implants (Papanicolaou et al., 2005). Finally, they can be placed flexibly according to prior hypotheses about the most likely seizure onset zone location, and adjusted based on the data recorded. This includes varying the sensor density based on need. At an engineering level, this would first require compatibility with movement, efficient noise cancellation, and ideally further miniaturization of the sensors (Sander et al., 2012).

A promising tendency more generally is the use of bigger datasets through sharing across labs and countries, as well as use of advanced mathematical models to tease out subtle data features which may be informative but hidden to experimenters and clinicians. The use of bigger datasets is presumed to be particularly promising in epilepsy where small and rare changes in the signal are easily overlooked but would likely be found using machine learning classification algorithms (Lecun et al., 2015). In addition to epilepsy, the use of more advanced data-driven mathematical models (from big datasets resulting from sharing of data across labs) could also be useful for sleep, psychiatry, pharmacology, and developmental neuroscience research. Further, such larger datasets would lend themselves well to neuro-feedback paradigms which in turn lend themselves well to several applications, e.g. in stroke recovery.

An important added benefit of the flexibility of OPM placement is the possibility of designing sensor arrays for children. At present, children and infants are rarely scanned in MEG due to the small head sizes which are not compatible with dewars optimized to fit a large proportion of the adult population. A small number of child MEG scanners exist but finding one which fits appropriate is unlikely given the rapid growth and consequent large range of head sizes of children. The possibility of using OPMs for child MEG is thus two-fold: not only would healthy brain processes and the variability hereof now be possible to measure and document, the ways in which development is retarded by specific diseases or disorders can be measured and tracked over time (from milliseconds to years) in spatiotemporal terms.

An obvious and non-trivial challenge with OPMs which are compatible with movement is how best to minimize the effect of movement or muscle-induced signal artefacts. One might imagine that neck and eye muscles for example would likely give rise to relatively large signals when a subject or patient is no longer asked or made to sit or lie still. Again, one might imagine that building advanced algorithms which for example were designed to make use of monitored muscular contractions (e.g. using electromyography) could eventually “learn” to account for movements. Perhaps this could even be done in conjunction with the Helmholtz-coil based field cancellation system if 3D tracking was used and compensated for online.

Proposed OPM head-cast design

An idea which emerged from the previous chapters was that if we can project empirically recorded hippocampal MEG activity back onto the scalp of a single subject, then we can begin to design and optimise the spatial configuration of OPM sensors for sampling this topography. In Chapter 5 we found that both location and orientation errors are detrimental to our ability to detect the hippocampus and must therefore be minimized *a priori*. One method for doing so is by constructing a solid head-cast which orients and fixes the sensors with respect to the head, effectively eliminating both independent and systematic errors in both location and orientation of the sensors. Such a design would allow experimenters to benefit from the high temporal resolution of MEG, the high signal amplitude attainable with OPMs, and the high spatial precision attainable with a head-cast.

Thus, to bring the experimental chapters together and provide an outlook for the future of hippocampal MEG, we translated the data from a single subject tested on the spatial memory task in Chapter 4 into a practical OPM-sensor array design (**Figure 6.1**).

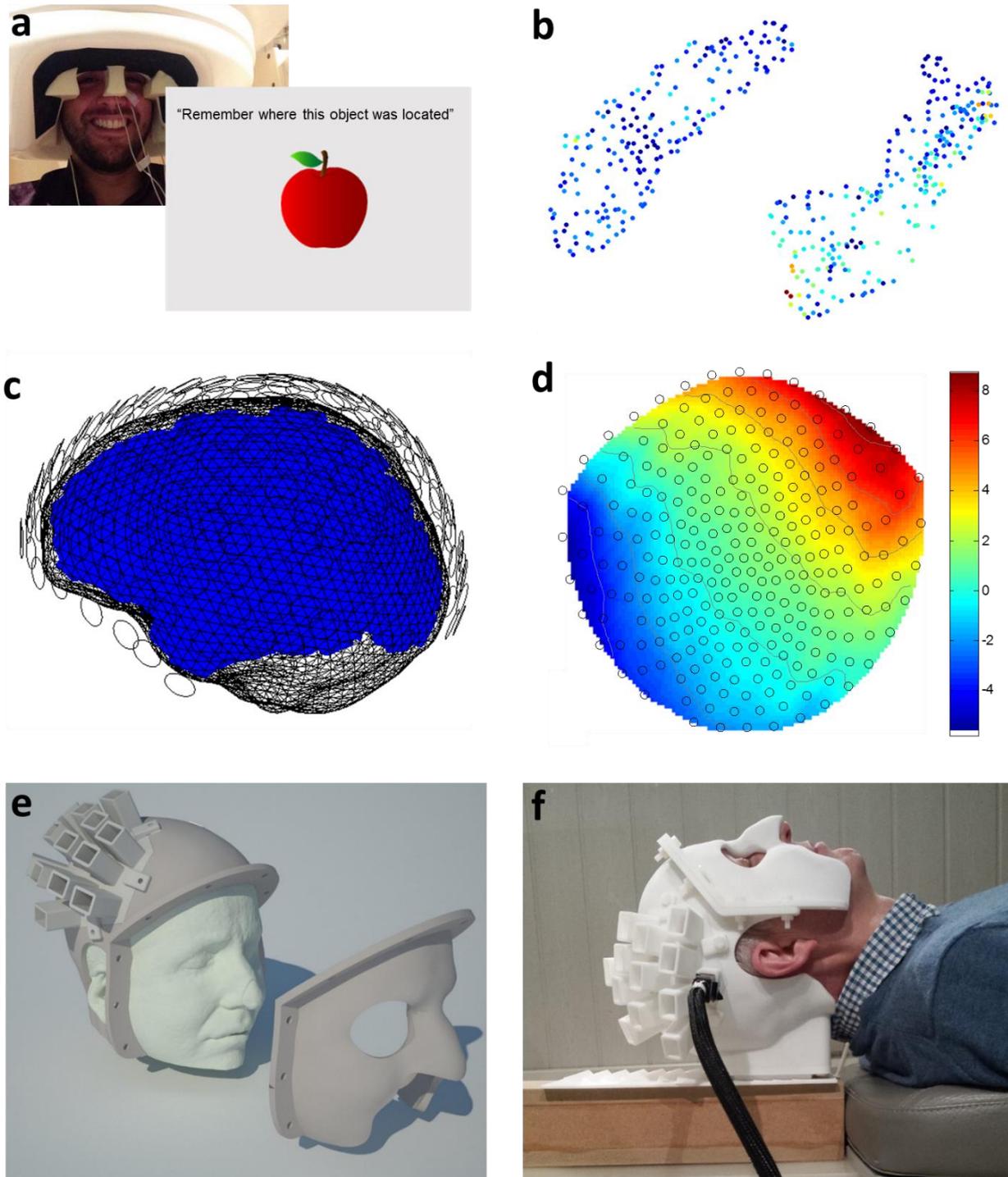


Figure 6.1 OPM sensor placement for optimal detection of hippocampus

a) Subject wearing head-cast was scanned in a normal MEG scanner and asked to perform a spatial memory experiment (described in Chapter 4). Data during spatial recall period was used: subjects were asked to remember the location of an object they had earlier “collected” in a virtual reality arena (insert illustration of one such object, an apple). The cued recall period was 3s long. **b)** Power in the 4-8Hz theta frequency band was calculated across all trials and all time. Figure shows normalized power values colour-coded for each hippocampal vertex.

The theta band peak for this subject is in the left anterior hippocampus. **c)** Spatial configuration of OPM sensors (black circles) placed 3 mm from the scalp of the same subject. Blue indicates the brain volume and the black mesh indicates the scalp shape. **d)** Scalp topography of the field pattern when projecting activity from the hippocampal peak onto the OPM sensor array. This 2D image suggests that optimal placement of sensors may be relatively widely distributed. Colour scale shows theta power intensity. **e)** Virtual 3D model of sensor placement for a median nerve experiment based on distribution of positive and negative portions of the field. The maximal density of sensors depends on sensor size. **f)** Photograph of same subject as used for hippocampal experiment, wearing N20-optimised 3D printed head-cast. Images e and f are adapted from Boto et al, 2016. (Although this placement and clustering of sensors is based on a different sensor topography and experiment from the present, the images bear similarity to how a hippocampal-specific 3D printed head-cast might look; in this case the sensors might be placed in two large clusters (as opposed to one) in order to optimally sample the positive and negative peaks found in d)).

The above figure shows the pipeline of the proposed head-cast design. Inspired by the design procedures outlined in Chapter 3 but with the location and orientation fixation constraint in mind, we first extracted the most active hippocampal vertex in the theta band across all trials for a single subject scanned on the spatial memory task in Chapter 4 (**Figure 6.1a,b**). We then plotted the scalp topography from the lead field of this peak vertex in a 2D plane, based on an OPM array with sensors placed 3 mm from the scalp surface (**Figure 6.1c,d**). This topography can then be used to determine how to sample the magnetic fields produced by this subject's hippocampus during this task. It shows relatively far apart maxima and minima of the field, which accords which patterns observed in concurrent MEG and intracranial data (Dalal et al., 2013a). If the number of sensors available were unlimited, it would therefore be optimal to sample the whole distribution (as in **Figure 6.1c**), but if not, one might choose to cluster sensors around the locations of the two extrema, and place some sensors where little if any field changes would be expected (i.e. in the green band spanning from left frontal to right posterior sensors, **Figure 6.1d**). It would also be possible to place sensors in or on parts of the face or neck which are not usually well-detected by a standard MEG array, but which may improve the sensitivity and spatial resolution of hippocampal sources even further. We are in the process of developing a 3D printed subject-specific head-cast optimised not for measuring hippocampus, but for a focal somatosensory evoked response from median nerve stimulation (**Figure 6.1e,f**). This design stabilizes the sensor orientations and locations in relation to the head (no systematic error) and perhaps more importantly, in relation to each other (no independent error). It could be

modified in terms of sensor placement and density based on the scalp topography from our hippocampal experiment and thereby enable hippocampal OPM-based MEG experiments to be conducted for the first time. It is the hope that in time, it will be possible to not only obtain dramatically improve hippocampal SNR data with MEG through the use of OPMs in this way, but also that subjects will be able to perform naturalistic tasks where they move freely in the environment, opening up new avenues of neuroimaging research in spatial navigation, social interaction, psychophysics, motor control et cetera.

We are hopeful that we have identified a suitable and relevant set of research questions to help further our understanding of possibilities and limitations for reliably detecting hippocampal signals using MEG. We designed an MEG simulation experiment and model (Chapter 1), recording device (Chapter 2), task adaptation and novel (rotation) analysis (Chapter 3), and OPM configuration (Discussion) for testing our hypotheses. The overall conclusion from the studies presented is that by carefully stabilizing and accurately measuring the spatial relationships between the sensors and the scalp, it becomes possible to improve the sensitivity to hippocampal sources in MEG. Future studies might build on this work either by using flexible head-casts with standard MEG scanners, or rigid head-casts with OPMs.

References

- Adjamian, P., Barnes, G.R., Hillebrand, a, Holliday, I.E., Singh, K.D., Furlong, P.L., Harrington, E., Barclay, C.W., Route, P.J.G., 2004. Co-registration of magnetoencephalography with magnetic resonance imaging using bite-bar-based fiducials and surface-matching. *Clin. Neurophysiol.* 115, 691–8. doi:10.1016/j.clinph.2003.10.023
- Alem, O., Benison, A.M., Barth, D.S., Kitching, J., Knappe, S., 2014. Magnetoencephalography of epilepsy with a microfabricated atomic magnetorode. *J. Neurosci.* 34, 14324–7. doi:10.1523/JNEUROSCI.3495-14.2014
- Allred, J.C., Lyman, R.N., Kornack, T.W., Romalis, M. V, 2002. High-sensitivity atomic magnetometer unaffected by spin-exchange relaxation. *Phys. Rev. Lett.* 89, 130801. doi:10.1103/PhysRevLett.89.130801
- Amaral, D.G., Witter, M.P., 1989. The three-dimensional organization of the hippocampal formation: A review of anatomical data. *Neuroscience* 31, 571–591. doi:10.1016/0306-4522(89)90424-7
- Andra, W., Nowak, H., 1998. *Magnetism in Medicine: A Handbook*. Wiley-VCH, Berlin.
- Attal, Y., Bhattacharjee, M., Yelnik, J., Cottureau, B., Lefèvre, J., Okada, Y., Bardinet, E., Chupin, M., Baillet, S., 2007. Modeling and detecting deep brain activity with MEG & EEG. *Conf. Proc. IEEE Eng. Med. Biol. Soc.* 2007, 4937–40. doi:10.1109/IEMBS.2007.4353448
- Attal, Y., Maess, B., Friederici, A., David, O., De, C., 2012. Head models and dynamic causal modeling of subcortical activity using magnetoencephalographic / electroencephalographic data 23, 85–95. doi:10.1515/RNS.2011.056
- Attal, Y., Schwartz, D., 2013. Assessment of Subcortical Source Localization Using Deep Brain Activity Imaging Model with Minimum Norm Operators: A MEG Study. *PLoS One* 8, e59856. doi:10.1371/journal.pone.0059856
- Backus, A.R., Schoffelen, J.-M., Szebényi, S., Hanslmayr, S., Doeller, C.F., 2016. Hippocampal-Prefrontal Theta Oscillations Support Memory Integration. *Curr. Biol.* 1–8. doi:10.1016/j.cub.2015.12.048
- Baillet, S., Mosher, J.C., Leahy, R.M., 2001. Electromagnetic brain mapping. *IEEE Signal Process. Mag.* 18, 14–30. doi:10.1109/79.962275
- Barron, H.C., Dolan, R.J., Behrens, T.E.J., 2013. Online evaluation of novel choices by simultaneous representation of multiple memories. *Nat. Neurosci.* 16, 1492–1498. doi:10.1038/nn.3515
- Barth, D.S., 2003. Submillisecond synchronization of fast electrical oscillations in neocortex. *J. Neurosci.* 23, 2502–10. doi:10.1523/JNEUROSCI.23/6/2502 [pii]

- Bastos, A.M., Usrey, W.M., Adams, R.A., Mangun, G.R., Fries, P., Friston, K.J., 2012. Canonical Microcircuits for Predictive Coding. *Neuron*. doi:10.1016/j.neuron.2012.10.038
- Baumgartner, C., Patarraia, E., Lindinger, G., Deecke, L., 2000. Neuromagnetic recordings in temporal lobe epilepsy. *J Clin Neurophysiol* 17, 177–189.
- Belardinelli, P., Jalava, a, Gross, J., Kujala, J., Salmelin, R., 2013. Optimal spatial filtering for brain oscillatory activity using the Relevance Vector Machine. *Cogn. Process.* 14, 357–69. doi:10.1007/s10339-013-0568-y
- Belardinelli, P., Ortiz, E., Barnes, G., Noppeney, U., Preissl, H., 2012. Source reconstruction accuracy of MEG and EEG Bayesian inversion approaches. *PLoS One* 7, e51985. doi:10.1371/journal.pone.0051985
- Bisby, J. a, King, J. a, Brewin, C.R., Burgess, N., Curran, H.V., 2010. Acute effects of alcohol on intrusive memory development and viewpoint dependence in spatial memory support a dual representation model. *Biol. Psychiatry* 68, 280–6. doi:10.1016/j.biopsych.2010.01.010
- Bisiach, E., Luzzatti, C., 1978. Unilateral neglect of representational space. *Cortex* 14, 129–133. doi:10.1016/S0010-9452(78)80016-1
- Bonnici, H.M., Chadwick, M.J., Kumaran, D., Hassabis, D., Weiskopf, N., Maguire, E. a, 2012. Multi-voxel pattern analysis in human hippocampal subfields. *Front. Hum. Neurosci.* 6, 290. doi:10.3389/fnhum.2012.00290
- Boto*, E., Meyer*, S., Shah, V., Alem, O., Knappe, S., Kruger, P., Fromhold, M., Lim, M., Morris, P.G., Bowtell, R., Barnes, G.R., Brookes, M.J., 2016. A New Generation of Magnetoencephalography: Room Temperature Measurements using Optically-Pumped Magnetometers. *Neuroimage* 1.
- Boto, E., Bowtell, R., Kruger, P., Fromhold, M.T., Morris, P.G., Meyer, S.S., Barnes, G.R., Brookes, M.J., 2016. On the Potential of a new generation of Magnetometers for MEG: a Beamformer Simulation Study. *PLoS One*. doi:10.1371/journal.pone.0157655
- Bragin, A., Engel, J., Wilson, C.L., Fried, I., Mathern, G.W., 1999. Hippocampal and entorhinal cortex high-frequency oscillations (100-500 Hz) in human epileptic brain and in kainic acid-treated rats with chronic seizures. *Epilepsia* 40, 127–137. doi:10.1111/j.1528-1157.1999.tb02065.x
- Bragin, A., Wilson, C.L., Engel, J., 2002. Rate of interictal events and spontaneous seizures in epileptic rats after electrical stimulation of hippocampus and its afferents. *Epilepsia* 43, 81–85. doi:10.1046/j.1528-1157.2002.35501.x
- Brambilla, P., Perlini, C., Rajagopalan, P., Saharan, P., Rambaldelli, G., Bellani, M., Dusi, N., Cerini, R., Pozzi Mucelli, R., Tansella, M., Thompson, P.M., 2013. Schizophrenia severity, social functioning and hippocampal neuroanatomy: three-dimensional mapping

- study. *Br. J. Psychiatry* 202, 50–5. doi:10.1192/bjp.bp.111.105700
- Brandon, M.P., Bogaard, A.R., Christopher, P., Michael, a, Gupta, K., Hasse, M.E., 2011. Periodicity from Directional Tuning. *Science* (80-.). 332, 595–599. doi:10.1126/science.1201652.Reduction
- Brebbia, C.A., Dominguez, J., 1989. *Boundary Elements: An Introductory Course*. McGraw-Hill, New York.
- Brookes, M.J., Vrba, J., Robinson, S.E., Stevenson, C.M., Peters, A.M., Barnes, G.R., Hillebrand, A., Morris, P.G., 2008. Optimising experimental design for MEG beamformer imaging. *Neuroimage* 39, 1788–1802. doi:10.1016/j.neuroimage.2007.09.050
- Brookes, M.J., Zumer, J.M., Stevenson, C.M., Hale, J.R., Barnes, G.R., Vrba, J., Morris, P.G., 2010. Investigating spatial specificity and data averaging in MEG. *Neuroimage* 49, 525–538. doi:10.1016/j.neuroimage.2009.07.043
- Burgess, N., 2008. Spatial cognition and the brain. *Ann. N. Y. Acad. Sci.* 1124, 77–97. doi:10.1196/annals.1440.002
- Burgess, N., Barry, C., O'Keefe, J., 2007. An oscillatory interference model of grid cell firing. *Hippocampus* 17, 801–812. doi:10.1002/hipo.20327
- Burgess, N., Maguire, E. a, O'Keefe, J., 2002. The human hippocampus and spatial and episodic memory. *Neuron* 35, 625–41.
- Burgess, N., O'Keefe, J., 2011. Models of place and grid cell firing and theta rhythmicity. *Curr. Opin. Neurobiol.* 21, 734–44. doi:10.1016/j.conb.2011.07.002
- Burke, J.F., Sharan, A.D., Sperling, M.R., Ramayya, A.G., Evans, J.J., Healey, M.K., Beck, E.N., Davis, K. a, Lucas, T.H., Kahana, M.J., 2014. Theta and high-frequency activity mark spontaneous recall of episodic memories. *J. Neurosci.* 34, 11355–65. doi:10.1523/JNEUROSCI.2654-13.2014
- Bush, D., Bisby, J., Bird, C., Gollwitzer, S., Rodinov, R., Scott, C., Diehl, B., McEvoy, A., Walker, M., Burgess, N., 2015. Hippocampal theta during memory guided virtual navigation in human intracranial EEG., in: *SfN Abstract*. p. 525.21.
- Buzsáki, G., 2005. Theta rhythm of navigation: link between path integration and landmark navigation, episodic and semantic memory. *Hippocampus* 15, 827–40. doi:10.1002/hipo.20113
- Buzsáki, G., Moser, E.I., 2013. Memory, navigation and theta rhythm in the hippocampal-entorhinal system. *Nat. Neurosci.* 16, 130–8. doi:10.1038/nn.3304
- Buzsáki, G.C.N.-K.S.U.O.R.O.U.A. 7Th F.Q.. B. 2006 C.. A.B.G.S.S.Q.. B. 2006 C.. D.U.E. 05-22-09 K.S.U.M.S.B.G.Q.. B. 2006 C.. A.W.S.F.S.W.L. 300, 2006. *Rhythms of the brain*. Oxford University Press, New York. doi:10.1093/acprof:oso/9780195301069.001.0001

- Byrne, P., Becker, S., 2004. Modeling mental navigation in scenes with multiple objects. *Neural Comput.* 16, 1851–72. doi:10.1162/0899766041336468
- Byrne, P., Becker, S., Burgess, N., 2007. Remembering the past and imagining the future: a neural model of spatial memory and imagery. *Psychol. Rev.* 114, 340–75. doi:10.1037/0033-295X.114.2.340
- Chupin, M., Baillet, S., Okada, Y.C., Hasboun, D., Garnero, L., 2002. On the Detection of Hippocampus Activity with {MEG}. 13th Int. Conf. Biomagn. August 8-14 2002 3.
- Cichy, R.M., Pantazis, D., 2015. Can visual information encoded in cortical columns be decoded from magnetoencephalography data in humans? *Arxiv Prepr.* 121, 1–44. doi:10.1016/j.neuroimage.2015.07.011
- Coffey, E.B.J., Herholz, S.C., Chepesiuk, A.M.P., Baillet, S., Zatorre, R.J., 2016. Cortical contributions to the auditory frequency-following response revealed by MEG. *Nat. Commun.* [In press], 1–11. doi:10.1038/ncomms11070
- Cohen, D., 1972. Magnetoencephalography: detection of the brain's electrical activity with a superconducting magnetometer. *Science* 175, 664–6. doi:10.1126/science.175.4022.664
- Cornwell, B.R., Arkin, N., Overstreet, C., Carver, F.W., Grillon, C., 2012. Distinct contributions of human hippocampal theta to spatial cognition and anxiety. *Hippocampus* 22, 1848–59. doi:10.1002/hipo.22019
- Cornwell, B.R., Johnson, L.L., Holroyd, T., Carver, F.W., Grillon, C., 2008. Human hippocampal and parahippocampal theta during goal-directed spatial navigation predicts performance on a virtual Morris water maze. *J. Neurosci.* 28, 5983–90. doi:10.1523/JNEUROSCI.5001-07.2008
- Crespo-García, M., Zeiller, M., Leupold, C., Kreiselmeyer, G., Rampp, S., Hamer, H.M., Dalal, S.S., 2016. Slow-theta power decreases during item-place encoding predict spatial accuracy of subsequent context recall. *Neuroimage.* doi:10.1016/j.neuroimage.2016.08.021
- Curio, G., 2000. Linking 600-Hz “spikelike” EEG/MEG wavelets (“sigma-bursts”) to cellular substrates: concepts and caveats. *J. Clin. Neurophysiol.* 17, 377–396. doi:10.1097/00004691-200007000-00004
- Dalal, S.S., Jerbi, K., Bertrand, O., Adam, C., Ducorps, a, Schwartz, D., Garnero, L., Baillet, S., Martinerie, J., Lachaux, J.P., 2013a. Evidence for MEG detection of hippocampus oscillations and cortical gamma-band activity from simultaneous intracranial EEG. *Epilepsy Behav.* 28, 310–311. doi:10.1016/j.yebeh.2012.04.032
- Dalal, S.S., Jerbi, K., Bertrand, O., Adam, C., Ducorps, a, Schwartz, D., Garnero, L., Baillet, S., Martinerie, J., Lachaux, J.P., 2013b. Evidence for MEG detection of hippocampus oscillations and cortical gamma-band activity from simultaneous intracranial EEG.

Epilepsy Behav. 28, 310–311. doi:10.1016/j.yebeh.2012.04.032

Dalal, S.S., Jerbi, K., Bertrand, O., Adam, C., Ducorps, a, Schwartz, D., Garnero, L., Baillet, S., Martinerie, J., Lachaux, J.P., 2013c. Evidence for MEG detection of hippocampus oscillations and cortical gamma-band activity from simultaneous intracranial EEG.

Epilepsy Behav. 28, 310–311. doi:10.1016/j.yebeh.2012.04.032

Dale, A.M., Sereno, M.I., 1993. Improved Localizadon of Cortical Activity by Combining EEG and MEG with MRI Cortical Surface Reconstruction: A Linear Approach. J. Cogn. Neurosci. 5, 162–176. doi:10.1162/jocn.1993.5.2.162

Dang, H.B., Maloof, a. C., Romalis, M. V., 2010. Ultrahigh sensitivity magnetic field and magnetization measurements with an atomic magnetometer. Appl. Phys. Lett. 97, 19–21. doi:10.1063/1.3491215

de Araújo, D.B., Baffa, O., Wakai, R.T., 2002. Theta oscillations and human navigation: a magnetoencephalography study. J. Cogn. Neurosci. 14, 70–8. doi:10.1162/089892902317205339

Doeller, C.F., King, J. a, Burgess, N., 2008. Parallel striatal and hippocampal systems for landmarks and boundaries in spatial memory. Proc. Natl. Acad. Sci. U. S. A. 105, 5915–20. doi:10.1073/pnas.0801489105

Dupont-Roc, J., Haroche, S., Cohen-Tannoudji, C., 1969. Detection of very weak magnetic fields (10– 9 gauss) by 87 Rb zero-field level crossing resonances. Phys. Lett. A 28, 638–639. doi:10.1016/0375-9601(69)90480-0

Duvernoy, H.M., 2005. The Human Hippocampus: Functional Anatomy, Vascularization and Serial Sections with MRI. Springer Science & Business Media.

Eichenbaum, H., 2015. The Hippocampus as a Cognitive Map ... of Social Space. Neuron 87, 9–11. doi:10.1016/j.neuron.2015.06.013

Ekstrom, A.D., Caplan, J.B., Ho, E., Shattuck, K., Fried, I., Kahana, M.J., 2005. Human hippocampal theta activity during virtual navigation. Hippocampus 15, 881–889. doi:10.1002/hipo.20109

Ekstrom, A.D., Kahana, M.J., Caplan, J.B., Fields, T.A., Isham, E.A., Newman, E.L., Fried, I., 2003. Cellular networks underlying human spatial navigation. Nature 425, 184–8. doi:10.1038/nature01964

Engels, M.M. a., Hillebrand, A., van der Flier, W.M., Stam, C.J., Scheltens, P., van Straaten, E.C.W., 2016. Slowing of Hippocampal Activity Correlates with Cognitive Decline in Early Onset Alzheimer’s Disease. An MEG Study with Virtual Electrodes. Front. Hum. Neurosci. 10, 1–13. doi:10.3389/fnhum.2016.00238

Fischl, B., 2012. FreeSurfer. Neuroimage 62, 774–781. doi:10.1016/j.neuroimage.2012.01.021

Fishbine, B., 2003. SQUID Magnetometry, Harnessing the Power of Tiny Magnetic Fields

[WWW Document]. URL file:///C:/Users/smeyer/Downloads/SQUID Magnetometry, Harnessing the Power of Tiny Magnetic Fields _ Spring 2003_ Los Alamos National Laboratory.html

- Friston, K., Harrison, L., Daunizeau, J., Kiebel, S., Phillips, C., Trujillo-Barreto, N., Henson, R., Flandin, G., Mattout, J., 2008a. Multiple sparse priors for the M/EEG inverse problem. *Neuroimage* 39, 1104–20. doi:10.1016/j.neuroimage.2007.09.048
- Friston, K., Harrison, L., Daunizeau, J., Kiebel, S., Phillips, C., Trujillo-Barreto, N., Henson, R., Flandin, G., Mattout, J., 2008b. Multiple sparse priors for the M/EEG inverse problem. *Neuroimage* 39, 1104–20. doi:10.1016/j.neuroimage.2007.09.048
- Friston, K., Mattout, J., Trujillo-Barreto, N., Ashburner, J., Penny, W., 2007. Variational free energy and the Laplace approximation, *NeuroImage*.
- Geselowitz, D.B., 1967. On bioelectric potentials in an inhomogeneous volume conductor. *Biophys. J.* 7, 1–11.
- Givens, B., 1995. Low doses of ethanol impair spatial working memory and reduce hippocampal theta activity. *Alcohol. Clin. Exp. Res.* 19, 763–767. doi:10.1111/j.1530-0277.1995.tb01580.x
- Gordon, B.A., Blazey, T., Benzinger, T.L.S., Head, D., 2013. Effects of aging and Alzheimer's disease along the longitudinal axis of the hippocampus. *J. Alzheimers. Dis.* 37, 41–50. doi:10.3233/JAD-130011
- Grech, R., Cassar, T., Muscat, J., Camilleri, K.P., Fabri, S.G., Zervakis, M., Xanthopoulos, P., Sakkalis, V., Vanrumste, B., 2008. Review on solving the inverse problem in EEG source analysis. *J Neuroeng Rehabil* 5, 25. doi:10.1186/1743-0003-5-25
- Green, J., Arduini, A., 1954. Hippocampal Electrical Activity in Arousal. *J. Neuro* 17, 533–57. doi:10.1016/0304-3940(96)12650-1
- Griffith, W.C., Knappe, S., Kitching, J., 2010. Femtotesla atomic magnetometry in a microfabricated vapor cell 18, 27167–27172.
- Grimault, S., Nolden, S., Lefebvre, C., Vachon, F., Hyde, K., Peretz, I., Zatorre, R., Robitaille, N., Jolicoeur, P., 2014. Brain activity is related to individual differences in the number of items stored in auditory short-term memory for pitch: evidence from magnetoencephalography. *Neuroimage* 94, 96–106. doi:10.1016/j.neuroimage.2014.03.020
- Gross, J., Baillet, S., Barnes, G.R., Henson, R.N., Hillebrand, A., Jensen, O., Jerbi, K., Litvak, V., Maess, B., Oostenveld, R., Parkkonen, L., Taylor, J.R., van Wassenhove, V., Wibral, M., Schoffelen, J.-M., 2013. Good practice for conducting and reporting MEG research. *Neuroimage* 65, 349–63. doi:10.1016/j.neuroimage.2012.10.001
- Gross, J., Timmermann, L., Kujala, J., Salmelin, R., Schnitzler, a., 2003. Properties of MEG tomographic maps obtained with spatial filtering. *Neuroimage* 19, 1329–1336.

doi:10.1016/S1053-8119(03)00101-0

- Guitart-Masip, M., Barnes, G.R., Horner, A., Bauer, M., Dolan, R.J., Duzel, E., 2013. Synchronization of medial temporal lobe and prefrontal rhythms in human decision making. *J. Neurosci.* 33, 442–51. doi:10.1523/JNEUROSCI.2573-12.2013
- Hafting, T., Fyhn, M., Molden, S., Moser, M., Moser, E.I., 2005. Microstructure of a spatial map in the entorhinal cortex. *Nature* 436, 801–806. doi:10.1038/nature03721
- Hallez, H., Vanrumste, B., Grech, R., Muscat, J., De Clercq, W., Vergult, A., D'Asseler, Y., Camilleri, K.P., Fabri, S.G., Van Huffel, S., Lemahieu, I., 2007. Review on solving the forward problem in EEG source analysis. *J. Neuroeng. Rehabil.* 4, 46. doi:10.1186/1743-0003-4-46
- Hämäläinen, M., Hari, R., Ilmoniemi, R.J., Knuutila, J., Lounasmaa, O. V., 1993. Magnetoencephalography—theory, instrumentation, and applications to noninvasive studies of the working human brain. *Rev. Mod. Phys.* 65, 413–497. doi:10.1103/RevModPhys.65.413
- Hanslmayr, S., Staresina, B.P., Bowman, H., 2016. Oscillations and Episodic Memory: Addressing the Synchronization/Desynchronization Conundrum. *Trends Neurosci.* 39, 16–25. doi:10.1016/j.tins.2015.11.004
- Happer, W., Tang, H., 1973. Spin-exchange shift and narrowing of magnetic resonance lines in optically pumped alkali vapors. *Phys. Rev. Lett.* 31, 273–276. doi:10.1103/PhysRevLett.31.273
- Harrison, L.M., Penny, W., Ashburner, J., Trujillo-Barreto, N., Friston, K.J., 2007. Diffusion-based spatial priors for imaging. *Neuroimage* 38, 677–695. doi:10.1016/j.neuroimage.2007.07.032
- Hartley, T., Maguire, E. a., Spiers, H.J., Burgess, N., 2003. The well-worn route and the path less traveled: Distinct neural bases of route following and wayfinding in humans. *Neuron* 37, 877–888. doi:10.1016/S0896-6273(03)00095-3
- Hauk, O., Wakeman, D.G., Henson, R., 2011. Comparison of noise-normalized minimum norm estimates for MEG analysis using multiple resolution metrics. *Neuroimage* 54, 1966–74. doi:10.1016/j.neuroimage.2010.09.053
- Hebb, D.O., 1949. *The Organization of Behaviour*. Wiley, New York.
- Helbling, S., Teki, S., Callaghan, M.F., Sedley, W., Mohammadi, S., Griffiths, T.D., Weiskopf, N., Barnes, G.R., 2015. Structure predicts function: Combining non-invasive electrophysiology with in-vivo histology. *Neuroimage* 108, 377–85. doi:10.1016/j.neuroimage.2014.12.030
- Henson, R.N., Mattout, J., Phillips, C., Friston, K.J., 2009. Selecting forward models for MEG source-reconstruction using model-evidence. *Neuroimage* 46, 168–76. doi:10.1016/j.neuroimage.2009.01.062

- Henson, R.N., Wakeman, D.G., Litvak, V., Friston, K.J., 2011. A Parametric Empirical Bayesian Framework for the EEG/MEG Inverse Problem: Generative Models for Multi-Subject and Multi-Modal Integration. *Front. Hum. Neurosci.* 5, 76. doi:10.3389/fnhum.2011.00076
- Hill, D.L.G., Maurer, C.R., Maciunas, R.J., Barwise, J. a., Fitzpatrick, J.M., Wang, M.Y., 1998. Measurement of intraoperative brain surface deformation under a craniotomy. *Neurosurgery* 43, 514–526. doi:10.1007/BFb0056187
- Hillebrand, A., Barnes, G.R., 2011. Practical constraints on estimation of source extent with MEG beamformers. *Neuroimage* 54, 2732–40. doi:10.1016/j.neuroimage.2010.10.036
- Hillebrand, A., Barnes, G.R., 2005. Beamformer Analysis of MEG Data. *Int. Rev. Neurobiol.* 68, 149–171. doi:10.1016/S0074-7742(05)68006-3
- Hillebrand, A., Barnes, G.R., 2003. The use of anatomical constraints with MEG beamformers. *Neuroimage* 20, 2302–2313. doi:10.1016/j.neuroimage.2003.07.031
- Hillebrand, a., Barnes, G.R., 2002. A Quantitative Assessment of the Sensitivity of Whole-Head MEG to Activity in the Adult Human Cortex. *Neuroimage* 16, 638–650. doi:10.1006/nimg.2002.1102
- Hillebrand, a., Nissen, I. a., Ris-Hilgersom, I., Sijsma, N.C.G., Ronner, H.E., van Dijk, B.W., Stam, C.J., 2016. Detecting epileptiform activity from deeper brain regions in spatially filtered MEG data. *Clin. Neurophysiol.* 127, 2766–2769. doi:10.1016/j.clinph.2016.05.272
- Hironaga, N., Hagiwara, K., Ogata, K., Hayamizu, M., Urakawa, T., Tobimatsu, S., 2014. Proposal for a new MEG-MRI co-registration: A 3D laser scanner system. *Clin. Neurophysiol.* 125, 2404–2412. doi:10.1016/j.clinph.2014.03.029
- Hostage, C. a., Roy Choudhury, K., Doraiswamy, P.M., Petrella, J.R., 2013. Dissecting the Gene Dose-Effects of the APOE ϵ 4 and ϵ 2 Alleles on Hippocampal Volumes in Aging and Alzheimer's Disease. *PLoS One* 8. doi:10.1371/journal.pone.0054483
- Iglói, K., Doeller, C.F., Berthoz, A., Rondi-reig, L., Burgess, N., 2010. Lateralized human hippocampal activity predicts navigation based on sequence or place memory. doi:10.1073/pnas.1004243107/- /DCSupplemental.www.pnas.org/cgi/doi/10.1073/pnas.1004243107
- Jacobs, J., 2014. Hippocampal theta oscillations are slower in humans than in rodents: implications for models of spatial navigation and memory. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 369, 20130304. doi:10.1098/rstb.2013.0304
- Jacobs, J., Korolev, I.O., Caplan, J.B., Ekstrom, A.D., Litt, B., Baltuch, G., Fried, I., Schulze-Bonhage, A., Madsen, J.R., Kahana, M.J., 2010. Right-lateralized brain oscillations in human spatial navigation. *J. Cogn. Neurosci.* 22, 824–836. doi:10.1162/jocn.2009.21240

- Jacobs, J., Weidemann, C.T., Miller, J.F., Solway, A., Burke, J.F., Wei, X.-X., Suthana, N., Sperling, M.R., Sharan, A.D., Fried, I., Kahana, M.J., 2013. Direct recordings of grid-like neuronal activity in human spatial navigation. *Nat. Neurosci.* 16, 1188–90. doi:10.1038/nn.3466
- Jafarpour, a, Horner, a J., Fuentemilla, L., Penny, W.D., Düzel, E., 2013. Decoding oscillatory representations and mechanisms in memory. *Neuropsychologia* 51, 772–80. doi:10.1016/j.neuropsychologia.2012.04.002
- Jeewajee, A., Lever, C., Burton, S., O’Keefe, J., Burgess, N., 2008. Environmental novelty is signaled by reduction of the hippocampal theta frequency. *Hippocampus* 18, 340–348. doi:10.1002/hipo.20394
- Jensen, O., Lisman, J.E., 2005. Hippocampal sequence-encoding driven by a cortical multi-item working memory buffer. *Trends Neurosci.* 28, 67–72. doi:10.1016/j.tins.2004.12.001
- John, T., Kiss, T., Lever, C., Érdi, P., 2014. Anxiolytic drugs and altered hippocampal theta rhythms: the quantitative systems pharmacological approach. *Network* 25, 20–37. doi:10.3109/0954898X.2013.880003
- Johnson, C., Schwindt, P.D.D., Weisend, M., 2010. Magnetoencephalography with a two-color pump-probe, fiber-coupled atomic magnetometer. *Appl. Phys. Lett.* 97, 13–16. doi:10.1063/1.3522648
- Johnson, C.N., Schwindt, P.D.D., Weisend, M., 2013. Multi-sensor magnetoencephalography with atomic magnetometers. *Phys. Med. Biol.* 58, 6065–77. doi:10.1088/0031-9155/58/17/6065
- Josephson, B.D., 1962. Possible new effects in superconductive tunnelling. *Phys. Lett.* 1, 251–253. doi:10.1016/0031-9163(62)91369-0
- Kahana, M.J., Seelig, D., Madsen, J.R., 2001. Theta returns. *Curr. Opin. Neurobiol.* 11, 739–44.
- Kaplan, R., Bush, D., Bonnefond, M., Bandettini, P. a, Barnes, G.R., Doeller, C.F., Burgess, N., 2014. Medial prefrontal theta phase coupling during spatial memory retrieval. *Hippocampus* 10, 1–10. doi:10.1002/hipo.22255
- Kaplan, R., Doeller, C.F., Barnes, G.R., Litvak, V., Düzel, E., Bandettini, P. a, Burgess, N., 2012a. Movement-related theta rhythm in humans: coordinating self-directed hippocampal learning. *PLoS Biol.* 10, e1001267. doi:10.1371/journal.pbio.1001267
- Kaplan, R., Doeller, C.F., Barnes, G.R., Litvak, V., Düzel, E., Bandettini, P. a, Burgess, N., 2012b. Movement-related theta rhythm in humans: coordinating self-directed hippocampal learning. *PLoS Biol.* 10, e1001267. doi:10.1371/journal.pbio.1001267
- Kiebel, S.J., Daunizeau, J., Phillips, C., Friston, K.J., 2008. Variational Bayesian inversion of the equivalent current dipole model in EEG/MEG. *Neuroimage* 39, 728–741.

doi:10.1016/j.neuroimage.2007.09.005

- Kim, K., Begus, S., Xia, H., Lee, S.K., Jazbinsek, V., Trontelj, Z., Romalis, M. V., 2014. Multi-channel atomic magnetometer for magnetoencephalography: A configuration study. *Neuroimage* 89, 143–151. doi:10.1016/j.neuroimage.2013.10.040
- King, J. a., Burgess, N., Hartley, T., Vargha-Khadem, F., O'Keefe, J., 2002. Human hippocampus and viewpoint dependence in spatial memory. *Hippocampus* 12, 811–820. doi:10.1002/hipo.10070
- Koessler, L., Cecchin, T., Caspary, O., Benhadid, a, Vespignani, H., Maillard, L., 2011. EEG-MRI co-registration and sensor labeling using a 3D laser scanner. *Ann. Biomed. Eng.* 39, 983–95. doi:10.1007/s10439-010-0230-0
- Kominis, I.K., Kornack, T.W., Allred, J.C., Romalis, M. V., 2003. A subfemtotesla multichannel atomic magnetometer. *Nature* 422, 596–599. doi:10.1038/nature01484
- Korczyński, A.D., Schachter, S.C., Brodie, M.J., Dalal, S.S., Engel, J., Guekht, A., Hecimovic, H., Jerbi, K., Kanner, A.M., Johannessen Landmark, C., Mares, P., Marusic, P., Meletti, S., Mula, M., Patsalos, P.N., Reuber, M., Ryvlin, P., Štillová, K., Tuchman, R., Rektor, I., 2013. Epilepsy, cognition, and neuropsychiatry (Epilepsy, Brain, and Mind, part 2). *Epilepsy Behav.* doi:10.1016/j.yebeh.2013.03.012
- Kreisman, N.R., Soliman, S., Gozal, D., 2000. Regional differences in hypoxic depolarization and swelling in hippocampal slices. *J. Neurophysiol.* 83, 1031–8.
- Kumar, S., Narayan, Y., Amell, T., 2003. Power spectra of sternocleidomastoids, splenius capitis, and upper trapezius in oblique exertions. *Spine J.* 3, 339–350. doi:10.1016/S1529-9430(03)00077-9
- Kumaran, D., Maguire, E.A., 2007. Match Mismatch Processes Underlie Human Hippocampal Responses to Associative Novelty. *J. Neurosci.* 27, 8517–8524. doi:10.1523/JNEUROSCI.1677-07.2007
- Kutas, M., Dale, A., 1997. Electrical and magnetic readings of mental functions, in: Rugg, M.D. (Ed.), *Cognitive Neuroscience*. MIT Press.
- Lecun, Y., Bengio, Y., Hinton, G., 2015. Deep learning. *Nature* 521, 436–444. doi:10.1038/nature14539
- Lega, B.C., Jacobs, J., Kahana, M., 2012. Human hippocampal theta oscillations and the formation of episodic memories. *Hippocampus* 22, 748–61. doi:10.1002/hipo.20937
- Liuzzi, L., Gascoyne, L.E., Tewarie, P.K., Barratt, E.L., Boto, E., Brookes, M.J., 2016. Optimising experimental design for MEG resting state functional connectivity measurement. *Neuroimage*. doi:10.1016/j.neuroimage.2016.11.064
- Logothetis, N.K., Eschenko, O., Murayama, Y., Augath, M., Steudel, T., Evrard, H.C., Besserve, M., Oeltermann, a, 2012. Hippocampal-cortical interaction during periods of subcortical silence. *Nature* 491, 547–53. doi:10.1038/nature11618

- Lopes da Silva, F., 2013. EEG and MEG: Relevance to neuroscience. *Neuron* 80, 1112–1128. doi:10.1016/j.neuron.2013.10.017
- López, J.D., Litvak, V., Espinosa, J.J., Friston, K., Barnes, G.R., 2014. Algorithmic procedures for Bayesian MEG/EEG source reconstruction in SPM, *NeuroImage*.
- López, J.D., Penny, W.D., Espinosa, J.J., Barnes, G.R., 2012. A general Bayesian treatment for MEG source reconstruction incorporating lead field uncertainty. *Neuroimage* 60, 1194–204. doi:10.1016/j.neuroimage.2012.01.077
- Lopez, J.D., Troebinger, L., Penny, W., Espinosa, J.J., Barnes, G.R., 2013. Cortical surface reconstruction based on MEG data and spherical harmonics. *Conf. Proc. IEEE Eng. Med. Biol. Soc. 2013*, 6449–52. doi:10.1109/EMBC.2013.6611031
- Maguire, E.A., Burgess, N., O'Keefe, J., 1999. Human spatial navigation: cognitive maps, sexual dimorphism, and neural substrates. *Curr Opin Neurobiol* 9, 171–177. doi:S0959-4388(99)80023-3 [pii]
- Mattout, J., Henson, R.N., Friston, K.J., 2007. Canonical source reconstruction for MEG. *Comput. Intell. Neurosci.* 2007. doi:10.1155/2007/67613
- Mattout, J., Phillips, C., Penny, W.D., Rugg, M.D., Friston, K.J., 2006. MEG source localization under multiple constraints: An extended Bayesian framework. *Neuroimage* 30, 753–767. doi:10.1016/j.neuroimage.2005.10.037
- McFarland, W., Teitelbaum, H., Hedges, E., 1975. Relationship between hippocampal theta activity and running speed in the rat. *J. Comp. Physiol. Psychol.* 88, 324–328.
- Medvedovsky, M., Nenonen, J., Koptelova, A., Butorina, A., Paetau, R., Makela, J., Ahonen, A., Simola, J., Gazit, T., Taulu, S., 2015. Virtual MEG Helmet: Computer Simulation of an Approach to Neuromagnetic Field Sampling. *IEEE J. Biomed. Heal. Informatics* 1–1. doi:10.1109/JBHI.2015.2392785
- Medvedovsky, M., Taulu, S., Bikmullina, R., Paetau, R., 2007. Artifact and head movement compensation in MEG. *Neurol. Neurophysiol. Neurosci.* 4, 4.
- Michalareas, G., Vezoli, J., van Pelt, S., Schoffelen, J.M., Kennedy, H., Fries, P., 2016. Alpha-Beta and Gamma Rhythms Subserve Feedback and Feedforward Influences among Human Visual Cortical Areas. *Neuron* 89, 384–397. doi:10.1016/j.neuron.2015.12.018
- Mikuni, N., Nagamine, T., Ikeda, a, Terada, K., Taki, W., Kimura, J., Kikuchi, H., Shibasaki, H., 1997. Simultaneous recording of epileptiform discharges by MEG and subdural electrodes in temporal lobe epilepsy. *Neuroimage* 5, 298–306. doi:10.1006/nimg.1997.0272
- Mills, T., Lalancette, M., Moses, S.N., Taylor, M.J., Quraan, M. a, 2012. Techniques for detection and localization of weak hippocampal and medial frontal sources using beamformers in MEG. *Brain Topogr.* 25, 248–63. doi:10.1007/s10548-012-0217-2

- Mitzdorf, U., 1985. Current source-density method and application in cat cerebral cortex: investigation of evoked potentials and EEG phenomena. *Physiol. Rev.* 65, 37–100.
- Morris, R.G., Garrud, P., Rawlins, J.N., O'Keefe, J., 1982. Place navigation impaired in rats with hippocampal lesions. *Nature*. doi:10.1038/297681a0
- Moser, E.I., Kropff, E., Moser, M.-B., 2008. Place cells, grid cells, and the brain's spatial representation system. *Annu. Rev. Neurosci.* 31, 69–89. doi:10.1146/annurev.neuro.31.061307.090723
- Moses, S.N., Hanlon, F.M., Ryan, J.D., 2011. Dynamic Imaging of Deep Brain Structures with MEG : Contributions to Understanding Human Memory, in: *Magnetoencephalography*. doi:10.5772/29133
- Mosher, J.C., Baillet, S., Leahy, R.M., 2003. Equivalence of linear approaches in bioelectromagnetic inverse solutions. *IEEE Work. Stat. Signal Process.* 2003 294–297. doi:10.1109/SSP.2003.1289402
- Murakami, S., Okada, Y., 2015. Invariance in current dipole moment density across brain structures and species: Physiological constraint for neuroimaging. *Neuroimage* 111, 49–58. doi:10.1016/j.neuroimage.2015.02.003
- Murakami, S., Okada, Y., 2006. Contributions of principal neocortical neurons to magnetoencephalography and electroencephalography signals. *J. Physiol.* 575, 925–36. doi:10.1113/jphysiol.2006.105379
- Muthukumaraswamy, S.D., 2013. High-frequency brain activity and muscle artifacts in MEG/EEG: a review and recommendations. *Front. Hum. Neurosci.* 7, 138. doi:10.3389/fnhum.2013.00138
- Myers, N.E., Rohenkohl, G., Wyart, V., Woolrich, M.W., Nobre, A.C., Stokes, M.G., 2015. Testing sensory evidence against mnemonic templates. *Elife* 4, 1–25. doi:10.7554/eLife.09000
- Nadel, L., Hoscheidt, S., Ryan, L.R., 2013. Spatial cognition and the hippocampus: the anterior-posterior axis. *J. Cogn. Neurosci.* 25, 22–8. doi:10.1162/jocn_a_00313
- Nenonen, J., Nurminen, J., Kičić, D., Bikmullina, R., Lioumis, P., Jousmäki, V., Taulu, S., Parkkonen, L., Putaala, M., Kähkönen, S., 2012. Validation of head movement correction and spatiotemporal signal space separation in magnetoencephalography. *Clin. Neurophysiol.* 123, 2180–91. doi:10.1016/j.clinph.2012.03.080
- Nolte, G., 2003. The magnetic lead field theorem in the quasi-static approximation and its use for magnetoencephalography forward calculation in realistic volume conductors. *Phys. Med. Biol.* 48, 3637–3652. doi:10.1088/0031-9155/48/22/002
- Nunez, P.L., Silberstein, R.B., 2000. On the relationship of synaptic activity to macroscopic measurements: does co-registration of EEG with fMRI make sense? *Brain Topogr.* 13, 79–96.

- O'Donnell, R.D., Berkhout, J., Adey, W.R., 1974. Contamination of scalp EEG spectrum during contraction of cranio-facial muscles. *Electroencephalogr. Clin. Neurophysiol.* 37, 145–151. doi:10.1016/0013-4694(74)90005-4
- O'Keefe, J., 2007. Hippocampal neurophysiology in the behaving animal, in: Andersen, P., Morris, R., Amara, D., Bliss, T., O'Keefe, J. (Eds.), *The Hippocampus Book*. Oxford Neuroscience, Oxford, pp. 458–475.
- O'Keefe, J., 1976. Place units in the hippocampus of the freely moving rat. *Exp. Neurol.* 51, 78–109. doi:10.1016/0014-4886(76)90055-8
- O'Keefe, J., Nadel, L., 1978. *The Hippocampus as a Cognitive Map*, Journal of Chemical Information and Modeling. Clarendon Press, Oxford.
- O'Keefe, J., Recce, M.L., 1993. Phase relationship between hippocampal place units and the EEG theta rhythm. *Hippocampus* 3, 317–330. doi:10.1002/hipo.450030307
- Osipova, D., Takashima, A., Oostenveld, R., Fernández, G., Maris, E., Jensen, O., 2006. Theta and gamma oscillations predict encoding and retrieval of declarative memory. *J. Neurosci.* 26, 7523–31. doi:10.1523/JNEUROSCI.1948-06.2006
- Packard, M.G., McGaugh, J.L., 1996. Inactivation of hippocampus or caudate nucleus with lidocaine differentially affects expression of place and response learning. *Neurobiol. Learn. Mem.* 65, 65–72. doi:10.1006/nlme.1996.0007
- Papadelis, C., Leonardelli, E., Staudt, M., Braun, C., 2012. Can magnetoencephalography track the afferent information flow along white matter thalamo-cortical fibers? *Neuroimage* 60, 1092–105. doi:10.1016/j.neuroimage.2012.01.054
- Papanicolaou, A.C., Pataraja, E., Billingsley-Marshall, R., Castillo, E.M., Wheless, J.W., Swank, P., Breier, J.I., Sarkari, S., Simos, P.G., 2005. Toward the substitution of invasive electroencephalography in epilepsy surgery. *J. Clin. Neurophysiol.* 22, 231–237. doi:10.1097/01.WNP.0000172255.62072.E8
- Parkkonen, L., Fujiki, N., Mäkelä, J.P., 2009. Sources of auditory brainstem responses revisited: Contribution by magnetoencephalography. *Hum. Brain Mapp.* 30, 1772–1782. doi:10.1002/hbm.20788
- Penny, W.D., 2012. Comparing dynamic causal models using AIC, BIC and free energy. *Neuroimage* 59, 319–330. doi:10.1016/j.neuroimage.2011.07.039
- Penny, W.D., Roberts, S.J., 1999. Bayesian neural networks for classification: How useful is the evidence framework? *Neural Networks* 12, 877–892. doi:10.1016/S0893-6080(99)00040-4
- Phillips, C., Mattout, J., Rugg, M.D., Maquet, P., Friston, K.J., 2005. An empirical Bayesian solution to the source reconstruction problem in EEG. *Neuroimage* 24, 997–1011. doi:10.1016/j.neuroimage.2004.10.030
- Phillips, C., Rugg, M.D., Friston, K.J., 2002. Anatomically informed basis functions for EEG

- source localization: combining functional and anatomical constraints. *Neuroimage* 16, 678–695. doi:10.1006/nimg.2002.1143
- Plonsey, R., 1977. Action Potential Sources and Their Volume Conductor Fields. *Proc. IEEE* 65, 601–611. doi:10.1109/PROC.1977.10539
- Poch, C., Fuentemilla, L., Barnes, G.R., Düzel, E., 2011. Hippocampal theta-phase modulation of replay correlates with configural-relational short-term memory performance. *J. Neurosci.* 31, 7038–42. doi:10.1523/JNEUROSCI.6305-10.2011
- Polyakov, O., 2003. A Femtotesla Optical Magnetometer Suitable for Applications In Multi-channel Magnetic Imaging 1–20.
- Pouget, a, Driver, J., 2000. Relating unilateral neglect to the neural coding of space. *Curr. Opin. Neurobiol.* 10, 242–9.
- Protzner, A.B., Valiante, T.A., Kovacevic, N., 2010. Hippocampal signal complexity in mesial temporal lobe epilepsy : a noisy brain is a healthy brain 289–297.
- Quraan, M. a, Moses, S.N., Hung, Y., Mills, T., Taylor, M.J., 2011. Detection and localization of hippocampal activity using beamformers with MEG: a detailed investigation using simulations and empirical data. *Hum. Brain Mapp.* 32, 812–27. doi:10.1002/hbm.21068
- Reuter, M., Schmansky, N.J., Rosas, H.D., Fischl, B., 2012. Within-subject template estimation for unbiased longitudinal image analysis. *Neuroimage* 61, 1402–1418. doi:10.1016/j.neuroimage.2012.02.084
- Rice, J.K., Rorden, C., Little, J.S., Parra, L.C., 2013. Subject position affects EEG magnitudes. *Neuroimage* 64, 476–84. doi:10.1016/j.neuroimage.2012.09.041
- Riggs, L., Moses, S.N., Bardouille, T., Herdman, A.T., Ross, B., Ryan, J.D., 2009. A complementary analytic approach to examining medial temporal lobe sources using magnetoencephalography. *Neuroimage* 45, 627–42. doi:10.1016/j.neuroimage.2008.11.018
- Rigoux, L., Stephan, K.E., Friston, K.J., Daunizeau, J., 2014. Bayesian model selection for group studies - revisited. *Neuroimage* 84, 971–85. doi:10.1016/j.neuroimage.2013.08.065
- Ross, B., Charron, R.E.M., Jamali, S., 2011. Realignment of magnetoencephalographic data for group analysis in the sensor domain. *J. Clin. Neurophysiol.* 28, 190–201. doi:10.1097/WNP.0b013e3182121843
- Ruddle, R. a., Payne, S.J., Jones, D.M., 1997. Navigating buildings in "desk-top" virtual environments: Experimental investigations using extended navigational experience. *J. Exp. Psychol. Appl.* 3, 143–159. doi:10.1037/1076-898X.3.2.143
- Rutishauser, U., Ross, I.B., Mamelak, A.N., Schuman, E.M., 2010. Human memory strength is predicted by theta-frequency phase-locking of single neurons. *Nature* 464, 903–7. doi:10.1038/nature08860

- Sander, T.H., Preusser, J., Mhaskar, R., Kitching, J., Trahms, L., Knappe, S., 2012. Magnetoencephalography with a chip-scale atomic magnetometer. *Biomed. Opt. Express* 3, 981. doi:10.1364/BOE.3.000981
- Sarvas, J., 1987. Basic mathematical and electromagnetic concepts of the biomagnetic inverse problem. *Phys. Med. Biol.* 32, 11–22. doi:10.1088/0031-9155/32/1/004
- Schultz, C., Engelhardt, M., 2014. Anatomy of the hippocampal formation. *Front. Neurol. Neurosci.* 34, 6–17. doi:10.1159/000360925
- Seager, M. a., Johnson, L.D., Chabot, E.S., Asaka, Y., Berry, S.D., 2002. Oscillatory brain states and learning: Impact of hippocampal theta-contingent training. *Pnas* 99, 1616–1620. doi:10.1073/pnas.032662099
- Sederberg, P.B., Kahana, M.J., Howard, M.W., Donner, E.J., Madsen, J.R., 2003. Theta and gamma oscillations during encoding predict subsequent recall. *J. Neurosci.* 23, 10809–10814. doi:23/34/10809 [pii]
- Serruya, M.D., Sederberg, P.B., Kahana, M.J., 2014. Power shifts track serial position and modulate encoding in human episodic memory. *Cereb. Cortex* 24, 403–13. doi:10.1093/cercor/bhs318
- Shah, V.K., Wakai, R.T., 2013. A compact, high performance atomic magnetometer for biomedical applications. *Phys. Med. Biol.* 58, 8153–61. doi:10.1088/0031-9155/58/22/8153
- Shepherd, G.M., 2003. *The Synaptic Organization of the Brain*. Oxford University Press.
- Shine, J.P., Valdes-Herrera, J.P., Hegarty, M., Wolbers, T., 2016. The Human Retrosplenial Cortex and Thalamus Code Head Direction in a Global Reference Frame. *J. Neurosci.* 36, 6371–6381. doi:10.1523/JNEUROSCI.1268-15.2016
- Singh, K.D., Holliday, I.E., Furlong, P.L., Harding, G.F.A., 1997. Evaluation of MRI-MEG / EEG simulation strategies using Monte Carlo 2.
- Spiers, H.J., Maguire, E.A., 2007. A navigational guidance system in the human brain. *Hippocampus* 17, 618–626. doi:10.1002/hipo.20298
- Spiers, H.J., Maguire, E. a., 2006. Thoughts, behaviour, and brain dynamics during navigation in the real world. *Neuroimage* 31, 1826–1840. doi:10.1016/j.neuroimage.2006.01.037
- Spruston, N., 2008. Pyramidal neurons: dendritic structure and synaptic integration. *Nat. Rev. Neurosci.* 9, 206–221. doi:10.1038/nrn2286
- Spruston, N., Jonas, P., Sakmann, B., 1995. Dendritic glutamate receptor channels in rat hippocampal CA3 and CA1 pyramidal neurons. *J. Physiology* 482, 325–352. doi:10.1113/jphysiol.1995.sp020521
- Squire, L.R., Stark, C.E.L., Clark, R.E., 2004. the Medial Temporal Lobe*. *Annu. Rev.*

- Neurosci. 27, 279–306. doi:10.1146/annurev.neuro.27.070203.144130
- Staresina, B.P., Michelmann, S., Bonnefond, M., Jensen, O., Axmacher, N., Fell, J., 2016. Hippocampal pattern completion is linked to gamma power increases and alpha power decreases during recollection. *Elife* 5, 1–18. doi:10.7554/eLife.17397.001
- Stefan, H., Hummel, C., Scheler, G., Genow, a., Druschky, K., Tilz, C., Kaltenhäuser, M., Hopfengärtner, R., Buchfelder, M., Romstöck, J., 2003. Magnetic brain source imaging of focal epileptic activity: A synopsis of 455 cases. *Brain* 126, 2396–2405. doi:10.1093/brain/awg239
- Stenroos, M., 2016. Integral equations and boundary-element solution for static potential in a general piece-wise homogeneous volume conductor c, 1–13.
- Stephan, K.E., Penny, W.D., Daunizeau, J., Moran, R.J., Friston, K.J., 2009. Bayesian model selection for group studies. *Neuroimage* 46, 1004–1017. doi:10.1016/j.neuroimage.2009.03.025
- Stephen, J.M., Ranken, D.M., Aine, C.J., Weisend, M.P., Shih, J.J., 2005. Differentiability of simulated MEG hippocampal, medial temporal and neocortical temporal epileptic spike activity. *J. Clin. Neurophysiol.* 22, 388–401. doi:10.1097/01.WNP.0000172141.26081.78
- Stevenson, C., Brookes, M., López, J.D., Troebinger, L., Mattout, J., Penny, W., Morris, P., Hillebrand, A., Henson, R., Barnes, G., 2014. Does function fit structure? A ground truth for non-invasive neuroimaging. *Neuroimage*. doi:10.1016/j.neuroimage.2014.02.033
- Stokes, M.G., Wolff, M.J., Spaak, E., 2015. Decoding Rich Spatial Information with High Temporal Resolution. *Trends Cogn. Sci.* 19, 636–638. doi:10.1016/j.tics.2015.08.016
- Stolk, A., Todorovic, A., Schoffelen, J.-M., Oostenveld, R., 2013. Online and offline tools for head movement compensation in MEG. *Neuroimage* 68, 39–48. doi:10.1016/j.neuroimage.2012.11.047
- Terrazas, A., Krause, M., Lipa, P., Gothard, K., Barnes, C., McNaughton, B.L., 2005. Self-Motion and the Hippocampal Spatial Metric. *J. Neurosci.* 25, 8085–8096. doi:10.1523/JNEUROSCI.0693-05.2005
- Tesche, C.D., Karhu, J., 2000. Theta oscillations index human hippocampal activation during a working memory task. *Proc. Natl. Acad. Sci. U. S. A.* 97, 919–24.
- Tiporlini, V., Alameh, K., 2013. High sensitivity optically pumped quantum magnetometer. *Sci. World J.* 2013. doi:10.1155/2013/858379
- Todd, N., Josephs, O., Callaghan, M.F., Lutti, A., Weiskopf, N., 2015. Prospective motion correction of 3D echo-planar imaging data for functional MRI using optical tracking. *Neuroimage* 113, 1–12. doi:10.1016/j.neuroimage.2015.03.013
- Tolman, E.C., 1948. Cognitive maps in rats and men. *Psychol. Rev.* 55, 189–208. doi:10.1037/h0061626

- Troebling, L., López, J.D., Lutti, A., Bestmann, S., Barnes, G., 2014a. Discrimination of cortical laminae using MEG. *Neuroimage*. doi:10.1016/j.neuroimage.2014.07.015
- Troebling, L., López, J.D., Lutti, A., Bradbury, D., Bestmann, S., Barnes, G., 2014b. High precision anatomy for MEG. *Neuroimage* 86, 583–91. doi:10.1016/j.neuroimage.2013.07.065
- Uutela, K., Taulu, S., Hämäläinen, M., 2001. Detecting and correcting for head movements in neuromagnetic measurements. *Neuroimage* 14, 1424–31. doi:10.1006/nimg.2001.0915
- van Pelt, S., Boomsma, D.I., Fries, P., 2012. Magnetoencephalography in twins reveals a strong genetic determination of the peak frequency of visually induced γ -band synchronization. *J. Neurosci.* 32, 3388–92. doi:10.1523/JNEUROSCI.5592-11.2012
- Van Veen, B.D., van Drongelen, W., Yuchtman, M., Suzuki, a, 1997. Localization of brain electrical activity via linearly constrained minimum variance spatial filtering. *IEEE Trans. Biomed. Eng.* 44, 867–880. doi:10.1109/10.623056
- Viard, A., Doeller, C.F., Hartley, T., Bird, C.M., Burgess, N., 2011. Anterior hippocampus and goal-directed spatial decision making. *J. Neurosci.* 31, 4613–21. doi:10.1523/JNEUROSCI.4640-10.2011
- Vrba, J., Robinson, S.E., 2001. Signal processing in magnetoencephalography. *Methods* 25, 249–71. doi:10.1006/meth.2001.1238
- Watrous, A.J., Fried, I., Ekstrom, A.D., Watrous, A.J., Fried, I., Ekstrom, A.D., 2011. Behavioral correlates of human hippocampal delta and theta oscillations during navigation Behavioral correlates of human hippocampal delta and theta oscillations during navigation. *J. Neurophysiol.* 1747–1755. doi:10.1152/jn.00921.2010
- Weiskopf, N., Suckling, J., Williams, G., Correia M., M.M., Inkster, B., Tait, R., Ooi, C., Bullmore T., E.T., Lutti, A., 2013. Quantitative multi-parameter mapping of R1, PD*, MT, and R2* at 3T: A multi-center validation. *Front. Neurosci.* 7, 1–11. doi:10.3389/fnins.2013.00095
- Wennberg, R., Valiante, T., Cheyne, D., 2011. EEG and MEG in mesial temporal lobe epilepsy: Where do the spikes really come from? *Clin. Neurophysiol.* 122, 1295–1313. doi:10.1016/j.clinph.2010.11.019
- Whalen, C., Maclin, E.L., Fabiani, M., Gratton, G., 2008. Validation of a method for coregistering scalp recording locations with 3D structural MR images. *Hum. Brain Mapp.* 29, 1288–301. doi:10.1002/hbm.20465
- Whitham, E.M., Pope, K.J., Fitzgibbon, S.P., Lewis, T., Clark, C.R., Loveless, S., Broberg, M., Wallace, A., DeLosAngeles, D., Lillie, P., Hardy, A., Fronsco, R., Pulbrook, A., Willoughby, J.O., 2007. Scalp electrical recording during paralysis: Quantitative evidence that EEG frequencies above 20 Hz are contaminated by EMG. *Clin. Neurophysiol.* 118, 1877–1888. doi:10.1016/j.clinph.2007.04.027

- Wibral, M., Singer, W., Aru, J., Uhlhaas, P.J., Stru, E., 2013. 2013_The Phase of Thalamic Alpha Activity Modulates Cortical Gamma-Band Activity Evidence from Resting-State MEG Recordings.pdf 33, 17827–17835. doi:10.1523/JNEUROSCI.5778-12.2013
- Wipf, D., Nagarajan, S., 2009. A unified Bayesian framework for MEG/EEG source imaging. *Neuroimage* 44, 947–66. doi:10.1016/j.neuroimage.2008.02.059
- Wisse, L.E.M., Gerritsen, L., Zwanenburg, J.J.M., Kuijf, H.J., Luijten, P.R., Biessels, G.J., Geerlings, M.I., 2012. Subfields of the hippocampal formation at 7T MRI: In vivo volumetric assessment. *Neuroimage* 61, 1043–1049. doi:10.1016/j.neuroimage.2012.03.023
- Xia, H., Baranga, A.B., Hoffman, D., Romalis, M. V, 2006. Magnetoencephalography with an atomic magnetometer 2–4. doi:10.1063/1.2392722
- Yang, Y., Kim, S., Kim, J.H., 2008. Ischemic evidence of transient global amnesia: Location of the lesion in the hippocampus. *J. Clin. Neurol.* 4, 59–66. doi:10.3988/jcn.2008.4.2.59
- Zhang, H., Jacobs, J., 2015. Traveling Theta Waves in the Human Hippocampus 35, 12477–12487. doi:10.1523/JNEUROSCI.5102-14.2015
- Zhang, J., Liu, W., Chen, H., Xia, H., Zhou, Z., Mei, S., Liu, Q., Li, Y., 2014. Multimodal neuroimaging in presurgical evaluation of drug-resistant epilepsy. *NeuroImage Clin.* 4, 35–44. doi:10.1016/j.nicl.2013.10.017