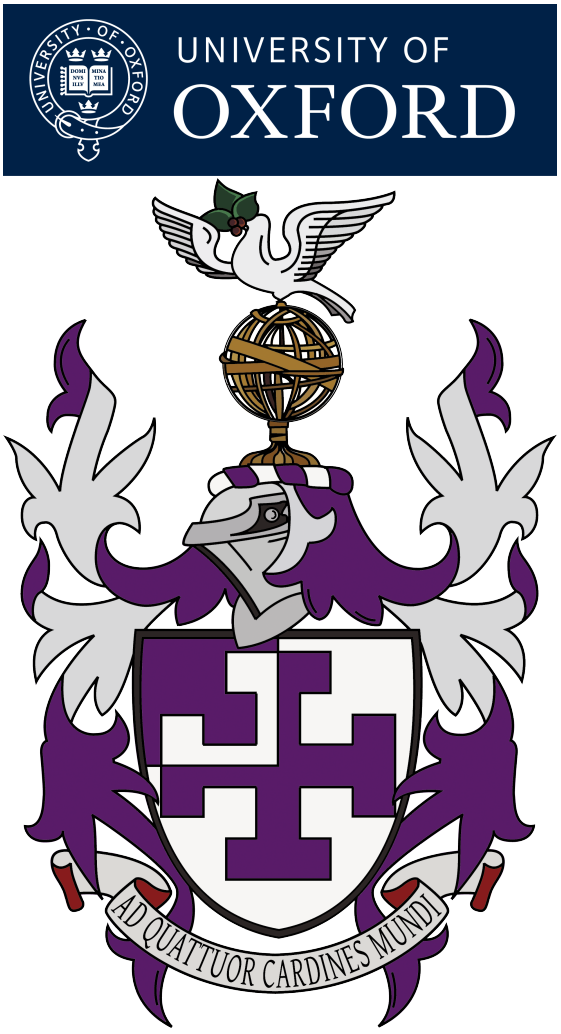
Time-varying Source Reconstruction

Ryan Charles Timms, MSci
St Cross College



A thesis submitted to the University of Oxford for the degree of Doctor of Philosophy

January 2022

Abstract

The human brain is a highly dynamical system. One way of imaging neural activity is through electrophysiological techniques such as magnetoencephalography and electroencephalography (M/EEG). However, for these data to be truly useful, the ill-posed inverse problem must first be tackled.

A fundamental step in solving this problem involves the introduction of assumptions about the data (and noise) covariance. To date, the majority of source localisation algorithms employ static approximations to the data covariance matrix.

Here we make the argument that researchers should consider a Temporally Adaptive SourcE Reconstruction (TASER) approach to solving the inverse problem. We achieve this by introducing a novel machine learning unsupervised recurrent neural network (RNN) based algorithm. Our model does not make restrictive assumptions about the temporal processes generating the data (i.e. it is not Markovian), nor does it assume that data originate from a set of mutually exclusive states. Instead, we model the data as a continuous mixture of temporally evolving *modes* or *elements*.

This algorithm, which we call VRAD, lets us add a temporal extension onto pre-existing M/EEG source reconstruction approaches. In this thesis we develop and explore the concepts of temporally adaptive spatial filters (beamformers) and generative model-based reconstruction procedures (e.g. multiple sparse priors).

This is motivated first by simulations which show an improvement in source reconstruction accuracy when a time-varying beamformer is used, and the ground truth source space data has a dynamic covariance profile. We then validate the algorithm on a previously published MEG head cast dataset Barratt et al. (2018). Our results show an improvement upon existing static techniques, and we are able to spatially distinguish activation on the order of a few millimetres between pinky and index finger movement. Despite never being exposed to the underlying task-structure, we show that VRAD is capable of learning physiologically plausible latent descriptions of the data. We further assess the philosophy of TASER on a more generic multi-modal multi-subject openly available M/EEG dataset published by Wakeman and Henson. We show that using multiple source reconstruction weights is not detrimental to the reconstruction of neural waveforms, and that dynamic techniques have benefits of static reconstruction algorithms.

Declaration

I confirm that the work presented in this thesis is my own, with the sole exception of Figure 3.7 which was authored by Dr Chetan Gohil. This is presented here with his permission.

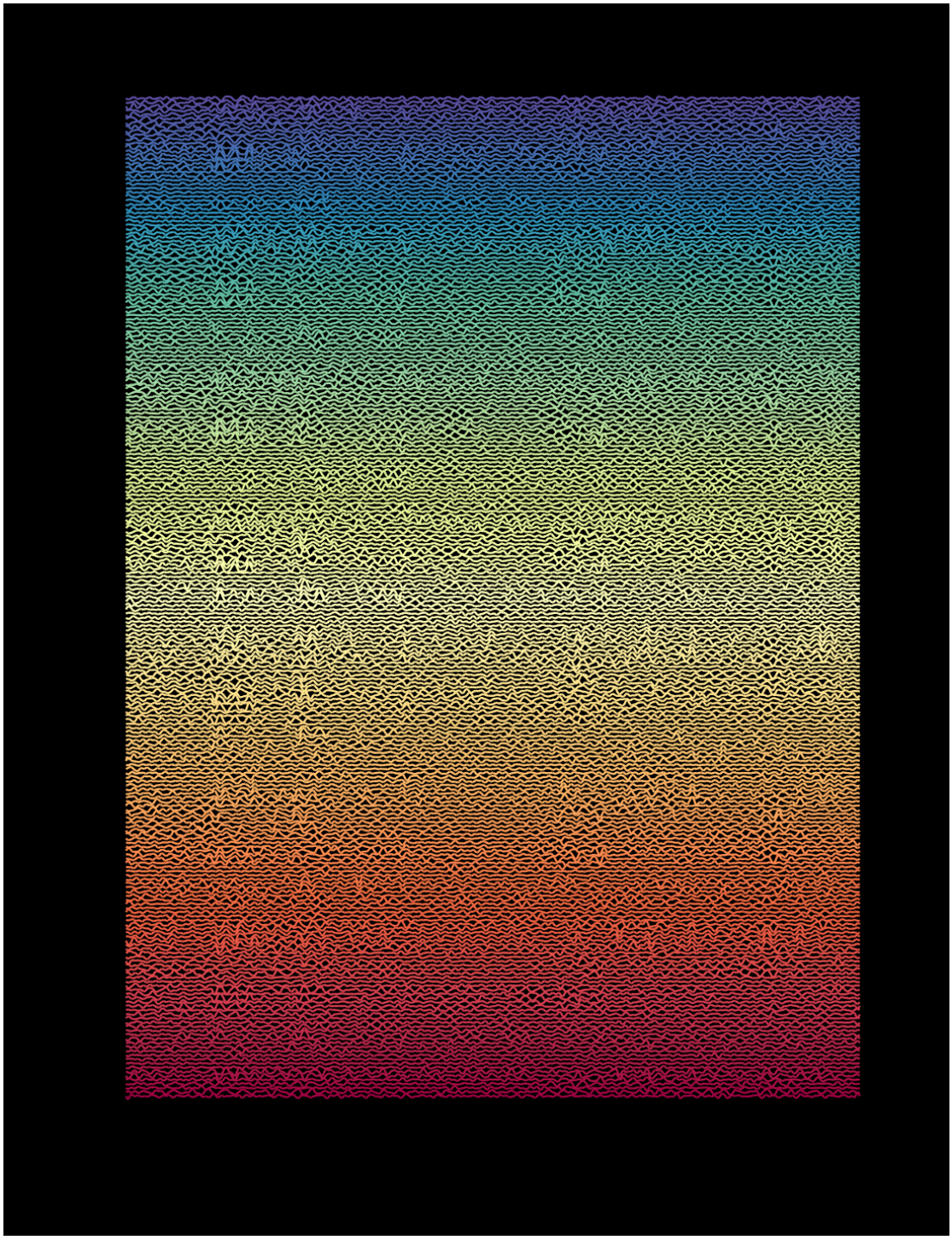
Signed:

Date:

"You truly belong here with us among the clouds"

Lando Calrissian, Star Wars

For my parents



Contents

Abstract . . . 1
Declaration . . . 2

List of Figures **5**

Acronyms **7**

Acknowledgements **8**

1 Introduction **10**
Non-Invasive Functional Brain Imaging . . . 11
Machine Learning . . . 12
Structure and Objectives of this Thesis . . . 14

2 Magnetoencephalography: from source to sensor, and back again **16**
Signal origins in MEG . . . 16
Neuron Fundamentals . . . 16
Action Potentials . . . 17
Neurotransmitters and post-synaptic currents . . . 18
The human brain . . . 19
Pyramidal neurons . . . 19
The biophysics of the recorded MEG signal . . . 19
Modern MEG systems . . . 21
Magnetic Sensors . . . 21
Magnetic Shielding . . . 22
Sensor Configurations . . . 22
The Inverse Problem: Definition . . . 22
Prerequisites in solving the inverse problem . . . 24
Coregistration . . . 24
Coordinate Systems . . . 24
Implications for source localisation . . . 24
The Forward Model . . . 25
The Single Sphere Head Model . . . 25
The Overlapping Spheres Head Model . . . 26
The Single Shell Head Model . . . 26
The Lead Field . . . 26
Formulism . . . 27
Source Orientation . . . 27
Existing Inversion Algorithms . . . 29
Equivalent Current Dipoles . . . 29
Beamformers . . . 29
Historic Applications of Beamformers . . . 30
Problems With Beamformers . . . 30
Covariance Estimation and Inversion . . . 30
Depth Bias . . . 31
Coherent Sources . . . 32
Tomographic Bayesian Approaches . . . 32
Reflection . . . 35
Conclusion . . . 35

3 A new approach to modelling **37**
Introduction to chapter . . . 37
Dynamics in the human brain . . . 37
Functional Magnetic Resonance Imaging Studies . . . 38
Electrophysiological Studies . . . 38
Existing approaches to modelling (dynamic) covariance . . . 38
A quick word on notation . . . 39
Sliding window analyses (SWAs) . . . 39
Drawbacks of the sliding window technique . . . 39
Hidden Markov Models . . . 40
Alternative HMMs . . . 41
Variational Recurrent Auto Decoder (VRAD) . . . 41
Generative Model . . . 42
Nomenclature . . . 43
Inference . . . 43
Amortised Inference . . . 43
Cost Function . . . 44
Definition . . . 44
The Reparameterisation Trick (RT) . . . 45
Recurrent Neural Networks . . . 45
Implementation . . . 46
Initialising The Inference . . . 47
Mini batch stochastic gradient descent . . . 48
KL Annealing . . . 48
Activation Functions . . . 48
Temperature . . . 49
Data Normalisation . . . 49
Simulations . . . 49
Showing the utility of a RNN-powered model . . . 49
VRAD can learn non-categorical descriptions of dynamic covariance . . . 51
Discussion . . . 52
Conclusion . . . 53

4 Temporally Adaptive SourcE Reconstruction (TASER) **54**
Introduction to chapter . . . 54
Previous non-stationary source reconstruction attempts . . . 54
Sliding Window Beamformers . . . 54
HMM Beamformers . . . 55
The TASER Approach . . . 55
Temporally adaptive multiple sparse priors (TAMSP) . . . 56
Temporally adaptive minimum norm estimate (TAMNE) . . . 57
Temporally adaptive beamformers (TABFER) . . . 57
Dimensionality Reduction . . . 57
Simulations . . . 58
Setup . . . 59
Results . . . 60
Simulation 1 . . . 60
Simulation 2 . . . 62
Simulation 3 . . . 62
Discussion . . . 63
Conclusion . . . 66

5 The Pinky and The Brain **67**
Introduction . . . 67
Background . . . 67
Dataset Synopsis . . . 68
Methods . . . 68
Ethical Consent . . . 69

Experimental Paradigm . . . 69
Foam Headcast Construction . . . 69
MEG Acquisition . . . 69
Preprocessing . . . 70
Analysis . . . 70
Statistical Procedure . . . 70
Beamforming . . . 71
Results . . . 73
Different temporal models give different perspectives on the same data 73
Incorporating data dynamics leads to more accurate source localisation 74
A TAMSP Result . . . 74
VRAD Beamformer Results . . . 74
HMM Beamformer Results . . . 74
Assessing Performance74
Learnt dynamics and spatial modes represent physiological processes . 77
Comparison with a HMM . . . 81
Gaining confidence in our results with state activations . . . 82
Dimensionality Reduction via K-means clustering . . . 83
Discussion . . . 83
Source Reconstruction Results . . . 83
Headcasts . . . 84
OPMs . . . 85
Accounting for forward model uncertainty with a generative model . . . 85
Beta bursts are identified by the inference . . . 85
Quantifying Covariance Estimation Uncertainty . . . 86
Oversimplifying our results . . . 86
Cross-modality priors . . . 86
Conclusion . . . 86
Acknowledgements . . . 87
Appendix 5.1: TAMSP Pre-Processing . . . 88
Appendix 5.2: The importance of covariance inversion regularisation . . . 88

6 A more challenging dataset: reconstructing deep cortical activity 93
Introduction . . . 93
Background . . . 93
Dataset Synopsis . . . 94
Previous analyses of this dataset . . . 94
Methods . . . 95
Ethical Consent . . . 95
Experimental Paradigm . . . 95
MEG Acquisition . . . 95
Structural Data . . . 95
Preprocessing . . . 96
Source Reconstruction . . . 96
Statistical Analysis . . . 97
Results . . . 98
VRAD gives a different image of dynamic covariance . . . 98
Source Reconstruction Improves the SNR of Evoked Responses . . . 98
Source Reconstruction Elucidates the Role of the Fusiform Gyrus in Face Processing . . . 98
TABFER Recovers the Classic M170 . . . 99
Latent Variables Correlate with Task Structure . . . 100
Discussion . . . 102
Conclusion . . . 106

7 Discussion and Conclusion **107**
Lessons from this thesis . . . 107
The struggle with TAMSP in its current implementation . . . 109
Some arguments against TASER . . . 110
Claim: using multiple weights obfuscates source space images . . . 110
On distorting reconstructed signals in time . . . 111
Concern: what if we derive erroneous covariance matrices? . . . 111
What VRAD gives us . . . 112
Hyperparameter Choices . . . 113
Choosing J: The number of modes . . . 113
Choosing K: The number of centroids . . . 113
VRAD vs the HMM . . . 114
Outlook . . . 115
Where next for M/EEG Source Reconstruction Research? . . . 115
Future applications of TASER . . . 116
Conclusion . . . 117

A Appendix A: Coregistration **118**

B Appendix B: MEG and EEG Forward Models **120**
Problem Statement . . . 120
Formulation . . . 120

C Appendix C: Deriving Beamformer Weights **124**

D Appendix D: MAP Equation Derivation **125**

List of Figures

1.1 A MEG pipeline . . . 12
1.2 A potential confound in source localisation pipelines . . . 13

2.1 Neuron fundamentals . . . 17
2.2 Signal generators in MEG . . . 20
2.3 MEG sensor profiles . . . 23
2.4 The effect of source orientation. . . . 27
2.5 Head models and lead field strengths . . . 31
2.6 The spatial filtering philosophy . . . 33
2.7 Source covariance priors distinguish different algorithms . . . 34

3.1 Hidden Markov Models within the context of MEG . . . 40
3.2 The generative process for the VRAD model . . . 42
3.3 The model and inference networks used in VRAD . . . 46
3.4 Recurrent neural network basics . . . 47
3.5 Initialisation of the VRAD inference . . . 47
3.6 The effects of temperature on the softmax function . . . 49
3.7 VRAD has longer temporal memory than a Hidden Markov Model . . . 50
3.8 VRAD can learn non-categorical descriptions of covariance. . . . 51

4.1 The TASER Approach . . . 56
4.2 Temporal data reduction . . . 58
4.3 Simulation setup for identifying dynamic reconstruction . . . 60
4.4 Simulation 1 Results: Correlations over SNRs . . . 61
4.5 Simulation 1 Results: Correlation change as a function of lead field strength . . . 62
4.6 Simulation 1 Results: Change in correlation is not a function of nearest neighbour distance . . . 63
4.7 Simulation 2 Results: Results generalise to real noise recordings . . . 63
4.8 Simulation 2 Results: An exemplar reconstructed timecourse . . . 64
4.9 Simulation 3 Results: Dataset length and dynamic reconstruction . . . 64

5.1 Motor experiment paradigm . . . 69
5.2 Analysis pipeline . . . 72
5.3 Gaining different perspectives of the same data . . . 73
5.4 A TAMSP Result . . . 75
5.5 Temporally adaptive beamformers outperform the static method in a precision dataset . . . 76
5.6 Across session results . . . 77
5.7 Inferred spatial modes . . . 78
5.8 Inferred mixing coefficients . . . 79
5.9 Beta bursts are found by VRAD . . . 80
5.10 Rescaling the VRAD coefficients . . . 81
5.11 HMM dynamics . . . 89
5.12 Latent variable timings give confidence to source localisation results . . . 90
5.13 Discretising inferred α_t values with K-means clustering . . . 91
5.14 Deriving VRAD statistics with a GMM . . . 92
5.15 The importance of eigenvalue truncation . . . 92

6.1 Face perception experimental paradigm 94
6.2 VRAD infers plausible latent descriptions of MEG data 99
6.3 Contrasting stationary and non-stationary beamformer results 100
6.4 VRAD stats at 168ms 101
6.5 TABFER recovers the classic M170 response in the right FFA 101
6.6 Non-stationary beamformers do not distort the classic M170 102
6.7 Spatial components learnt by unsupervised learning algorithms 103
6.8 Spatial and temporal elements correlate with task structure 104
6.9 Eigenspectra for the different covariances 105

7.1 A potential argument against TASER 112
7.2 VRAD gives a more dynamic description of data 116

Acronyms

SNR	Signal-to-Noise Ratio
MNE	Minimum Norm Estimate
MSP	Multiple Sparse Priors
MEG	Magnetoencephalography
EEG	Electroencephalography
LSTM	Long Short Term Memory
RNN	Recurrent Neural Network
VRAD	Variational Recurrent Auto Decoder
SGVB	Stochastic Gradient Variational Bayes
rsMEG	Resting State MEG
fMRI	Functional Magnetic Resonance Imaging
ECD	Equivalent Current Dipole
LCMV	Linearly Constrained Minimum Variance
EBB	Empirical Based Beamformer
MAP	Maximum a posteriori
PCA	Principal Component Analysis
ERN	Empty Room Noise
ROI	Region of Interest
ICA	Independent Component Analysis
HMM	Hidden Markov Model
VFE	Variational Free Energy
BPF	Band-Pass Filtered
MVN	Multivariate Normal

Acknowledgements

To name and shame her, my friend Cleo once likened a PhD to having a baby. There are moments of excitement prior to starting; a shock of the reality as you come to terms with things once you get under way and then a long, hard period of gestation over many months. Finally, there is the gruelling nature of bringing your new creation into the world. And before you know it, it's over. Poof. Done. All of the pain is forgotten and a huge surge of endorphins lead you to proudly proclaim that you would do it all over again. Well, some parents do. Others stick with one child - and I think one PhD is good enough for me.

Although I haven't been counting, I am sure that some combination of the words "thank", "you", "very" and "much" must have passed my lips well over 5,000 times over the past 4 years. I am unlikely to ever win at the academy awards, so please allow me some verbosity in this section, and to repeat such expressions of gratitude for the final time as a doctoral candidate.

To repeat the age-old and well-trodden cliché , this DPhil[1] thesis wouldn't ever have come to fruition if it wasn't for a huge investment of love, time and effort from family, teachers, mentors and, above all, friends both new and old, ever present and far away. My sincere apologies go out to anyone that I have omitted (accidentally, I promise).

The first person that I have to thank is my dear principal supervisor, Mark 'Wooly' Woolrich. Thank you for always having your office door open and for finding the time to talk (sometimes for hours on end!). Thank you for rekindling my inspiration at times and equally for dampening my hasty periods of excitement and keeping me on the right track when necessary. You have been a frighteningly bright source of new ideas throughout this degree and your creativity is something to behold. Above all, you have been a brilliant mentor and provided unwavering personal support and warmth. Thank you.

There truly is only one Dr Andrew Quinn. First as a friend and colleague, and later as a supervisor, you never failed to provide me with the best advice that a naïve student could have hoped for. I doubt that my academic debt to you can ever be repaid, but I will try. I owe you one. Bigly.

Thank you to Steve Smith for always being a friendly face and thought-provoking mentor, and for fostering such an incredible environment in Oxford.

This thesis is the culmination of 20 years of formal education, and I owe an enormous amount to some amazing teachers. Of note, special mentions must go to Professors Boucher, Paradis, Gordon, Boothroyd and Waters. Thank you all, and thanks to Wallington County Grammar School. An early mention here to Uncle Mark for all of the additional maths tuition during my A-levels and for getting me into University - I really, really wouldn't be here if it wasn't for all of those wintery Tuesday evenings in your kitchen. Thank you also to Jay Howells and Phil Gibson for their Sunday mornings growing up. Thank you to everyone in "Don't scroll up".

I look back at my time at The University of Nottingham with the fondest of memories. My undergraduate experience wouldn't have been half as productive, let alone fun, without all of you wonderful people. In no particular order, colossal thanks to James Thurgood, Ilaria Lisi, Zelekha Abid, Shane De Silva, Ryan Chapman, Dave Vernon, Alice Hibbert, and Louise Wells for filling my days with joy. Thank you Jackson Allen for always being there and for dragging me away from my work to take me to the Organ Grinder, and for the poker insights. Everyone at University of Nottingham

[1] Ah. Oxford speak on page 1. Much like I've been avoiding this for the past 4 years, I'll try and avoid this in the following pages. A 'DPhil' is just an Oxford abbreviation for a Doctor of Philosophy, or PhD.

Climbing Club whom I don’t have the space to name personally - our days out on crags around the most beautiful parts of the UK have kept me sane, but more importantly have often ended up with me having a stitch from laughing too hard and with tears rolling down my cheek. You all hold a special place in my heart. Special thanks to the Côte d’Ivoire crew - Ollie Ross, Alkis Tsapanidis and Astrid Southam.

Throughout my undergraduate days I was incredibly fortunate to get exposed to world class research, and given opportunities that most PhD candidates could only dream of. Such experiences were incredibly formative. Some previous supervisors and lecturers that I would like to name here are Matt Brookes, Karen Mullinger, George O’Neill, David Carmichael, Tim Tierney, Ellie Shamshiri, Jessica Maclean, Mark Fromhold, Liam Smith, Lucy Smith and James Sharp. Thank you for all of your hard work spent on teaching me and my peers, and for helping us on our way.

Thank you to the MEG community for being such a friendly and welcoming bunch of people. Special thanks to Jimmy Bonaiuto, Vladimir Litvak, Francesca Castaldo and Gareth Barnes. To my distant friends scattered around the globe who I’ve met at conferences, summer schools and whilst travelling, thank you for an amazing time over the years. See you soon.

Life in Oxford would have been a lot tougher if it wasn’t for my friends at the Oxford Centre for Human Brain Activity (OHBA); Giedre, Cameron Higgins, Diego Vidaurre, Jen Chesters, Catharina Zich (the best desk neighbour), Sage Boettcher, Dejan Draschkow, Evan Roberts, Jiewon Kang, Saed Khawaldeh, Lauren Atkinson, Usama, Jemma, Sarah, Girl Dani, Jasper, Richard and my brother from another mother, Mats. Special mentions must again go to Alex Skates and Andrew Quinn for getting me off the ground in the early days. Thanks for the laughs Michael, Isabelle, Ivor, Jacqueline, Amy, Irene, William Rooney, Gabriel, Alice and Tom Jones. Cheers to the guys at the Star and Marios, too!

Thank you to Dr Valentine, John Shiels, Karen and Dave. Special thanks to Peter Jezzard for always being of immense help.

Mum, Dad, Erin, Alex, Toby (the dog, RIP), Uncle Mark, Oliver, Dominic, Ella, Finlay, Archie, Aunty Tib, Uncle (Dr!) Preston, Granddad, Nanny Anne and Granddad Peter, Grandma, Uncle Simon, Jessica, Beth, Liam, Lucy, Fin, Rosalie, Aunty Clare, Adam, Steven and Becky you have never failed to support me in my endeavours. This thesis (and its typos!) is just as much yours as it is mine.

Finally, thanks to Nanny-noo for always giving me the best advice and for knowing what to do even in the trickiest of situations.

Ryan Timms
A basement in London, 2022.



1 Introduction

By the time you have finished reading this sentence, some 85 billion neurons in your brain will have made approximately 425 billion computations (Lent et al., 2012)[1]. Clearly, a lot is going on inside your head. Remarkably, during that time, your brain was able to control your eyes via the motor cortex, scanning them left-to-right across the page, before dropping them down a line to finish reading the sentence, all the meanwhile transferring a visual input from the retina to your primary visual cortex before relaying this information to temporal lobe areas, passing through Wernicke's area along the way. Subconsciously, the computer inside you simultaneously kept your heart beating and your lungs functioning. All of this language processing happened sequentially and on the order of a few seconds; your brain was (hopefully) able to make sense of the individual words before integrating their attributed definitions to make up a comprehensible narrative.

What can we deduce from these ideas? First, it appears that the brain comprises loosely distributed regions which have strong functional correlations with particular tasks; the motor cortex controls the muscles around our eyes, the visual cortex deals with sensory input from the eyes and Wernicke's area is associated with language. Moreover, these regions must somehow be communicating with one another. Second, as well as being highly intelligent conscious beings, our brain has a subconscious house-keeping element to it which keeps us alive. Third, and crucially for our well-being, operations in our brain are being carried out on at least a sub-second timescale.

Now consider the scenario that you're at an orchestral concert with a friend about to see a symphony. You arrive 30 minutes before the start of the show and take your seats in the auditorium. Like everyone else around you, you are contently chatting away and catching up with each other and talking through your respective days at work. On the stage, the curtain is yet to drop, and so sounds from the crowd are transformed as they pass through the fabric of the curtain.

Behind the curtain, minutes before the start of the performance, a sound engineer walks around the stage and turns on some 300 microphones for the rather elaborate group. Tonight's gig is being recorded for a live CD. For house-keeping purposes, the microphones start recording five minutes before the curtain drops, and so pick

[1] This is based on the approximation that there are 86 ± 8 billion neurons in the brain of a healthy male between the ages of 50 and 70, the assumption that each one of these neurons "spikes" about once every second (see: https://aiimpacts.org/rate-of-neuron-firing/) and that it took you roughly 5 seconds to read the sentence. Inspection of both references here shows that this is a hard thing to quantify, but isn't a terrible order of magnitude approximation.

up the amalgamation of conversations from the 10,000 person audience. Is there a way of somehow disentangling all of these sounds and catching up on the latest gossip from Sarah and Bradley who are sat in seats H1 and H2?

This is very analogous to the problem that we often want to solve when analysing magnetoencephalography (MEG) data - to deduce where in the brain activity is occurring, and what this signal is. In MEG, we non-invasively record the magnetic fields associated with the synchronous firing of tens of thousands of neurons from within the brain, with (typically) hundreds of sensors placed around the head. Like a microphone, the sampling rate of MEG is on the order of kHz, and so we obtain a measure of neuronal activity at timescales fundamental to the computations underlying cognition. Similar to the effect of the curtain in the concert example, with a good understanding of electromagnetism and classical physics, we are able to write down equations which describe how an electrical source in the brain produces a magnetic field, and how the anatomy of a person alters this signal before it is recorded by the sensor array. The neuroscientist now asks if there is a way of getting back to the original cortical signal from the measurements made outside of the head, and what are areas V1 and V2 of the visual cortex saying to each other?

Non-Invasive Functional Brain Imaging

Non-invasive functional neuroimaging is a key tool in trying to understand how the healthy brain functions in normal day-to-day life. Whilst invasive techniques such as electrocorticography (ECoG) and positron emission tomography (PET) can provide a more detailed (and often cleaner) image of neural activity, their applications are limited to diseased cohorts, limiting their utility in the general healthy population.

Functional magnetic resonance imaging (fMRI) has exploded in the past three decades and has led the charge in whole-brain imaging research. fMRI (and MRI, more generally) works by manipulating and recording the change in magnetisation of soft tissues. As these changes are acquired in K space, the inverse problem to reconstruct these signals into images is well-posed. This means that the spatial resolution of the technology is often, and reliably, exquisite.

This detail does come at a cost, however: the more refined we want our images to be, the more time we have to invest in sampling K space. This makes fMRI a fundamentally slow neuroimaging device, with sampling frequencies on the order of a second. More fundamentally, the blood oxygen level dependent (BOLD) signal detected in fMRI is only a secondary metabolic measure of true underlying electrical brain activity. This fluctuates on the scale of seconds, orders of magnitude slower than the true underlying timescales of cortical activity.

It is well known that neurons rely on chemical and electrical processes to communicate. Remarkably, when this communication is sufficiently synchronised in time and space, the resulting electromagnetic fields of these processes become large enough to be detected outside of the brain. This was first achieved famously by Berger (1929) in the case of electroencephalography (EEG) and Cohen (1968) for magnetoencephalography (MEG).

Electrophysiological provide unrivalled temporal detail on electrical activity originating from the brain. MEG and EEG are used to record the spontaneously occurring neural oscillations which we observe outside of the head. These range from the ultraslow, “delta waves" ($0 - 3.5Hz$), observed in sleep (Borbély et al., 1981) and stroke patients (Tecchio et al., 2007), to the ultrafast, “gamma waves" ($30Hz+$), involved in perception, cognition and general higher order processes. Between these extremes lie the theta band ($4 - 8Hz$), integral to goal-directed behaviour such as spatial navigation (Cornwell et al., 2008) and working memory in the hippocampus (Tesche

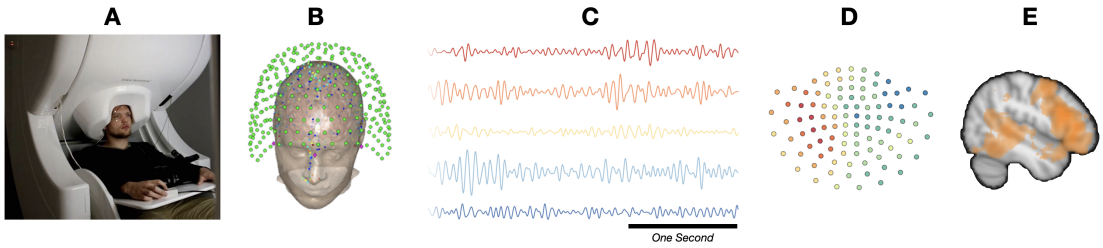


Figure 1.1: A stereotypical MEG study pipeline. A) Participants perform a task or just rest next to the liquid helium dewar whilst we non-invasively and silently record their brain activity. B) Functional data are imported and analysed in software packages alongside structural data. C) MEG offers sublime temporal resolution at the sensor level and can image both spontaneous and task-induced neural rhythms. D) These data can also be analysed across space at the MEG sensors, but more is gleaned when data are source localised to cortex as shown in E). This involves solving the famously ill-posed inverse problem.

and Karhu, 2000), and the alpha band ($8-13Hz$), the rhythm famously shown to be modulated in the visual cortex by Berger. Motor activity is commonly associated with brain waves in the beta band ($13-30Hz$) and is most prevalent over Rolandic areas.

MEG data are acquired by tens or hundreds of sensors which passively “listen" to what the brain is saying at each instance in time. Whilst grateful for being able to acquire these signals, the ambitious neuroscientist typically would like to know what every brain region (thousands of locations) is doing throughout an experiment, rather than merely what each sensor is picking up. This is the essence of *the M/EEG inverse problem* and it is *ill-posed*. Nevertheless, when efforts *are* made to solve this technically impossible challenge, MEG and EEG are unique in their ability to offer rich movies of neural activity at cognitive timescales. Solving this problem, of mapping 300 seemingly abstract timeseries recorded outside of the head into statistical images of neural power, “source reconstruction", is the primary goal of this thesis. Figure 1.1 combines these ideas and depicts a typical cutting-edge MEG pipeline.

The aforementioned temporal resolution boasted by MEG and EEG is fully exploited in certain domains of the field (e.g. time-frequency analyses). However, it is often overlooked or over-simplified in the source localisation community. As we will see in the proceeding chapters, source inversion algorithms centre around the calculation of statistics of the observed data (or making assumptions about them). For the most part, it is assumed either implicitly or explicitly that this statistic is best approximated by a single time-averaged value. This idea is shown in Figure 1.2. Perhaps this is done for the sakes of simplicity, yet making this assumption of stationarity is by no means a necessity. We will explore this idea in detail in this thesis, and ask whether this assumption of stationarity in time can lead to detrimental effects in the spatial domain when reconstructing data.

Solving the inverse problem in MEG necessitates the modelling of data. One such way of doing this modelling is through *machine learning*.

Machine Learning

Machine learning (ML) has enjoyed something of a renaissance over the past decade, with its popularity, uptake and fan-base exploding. Far from a new idea, improvements in graphical processing units (GPUs) and ever-increasing accessibility to large datasets has led to the technology becoming ubiquitous and underpinning our every day

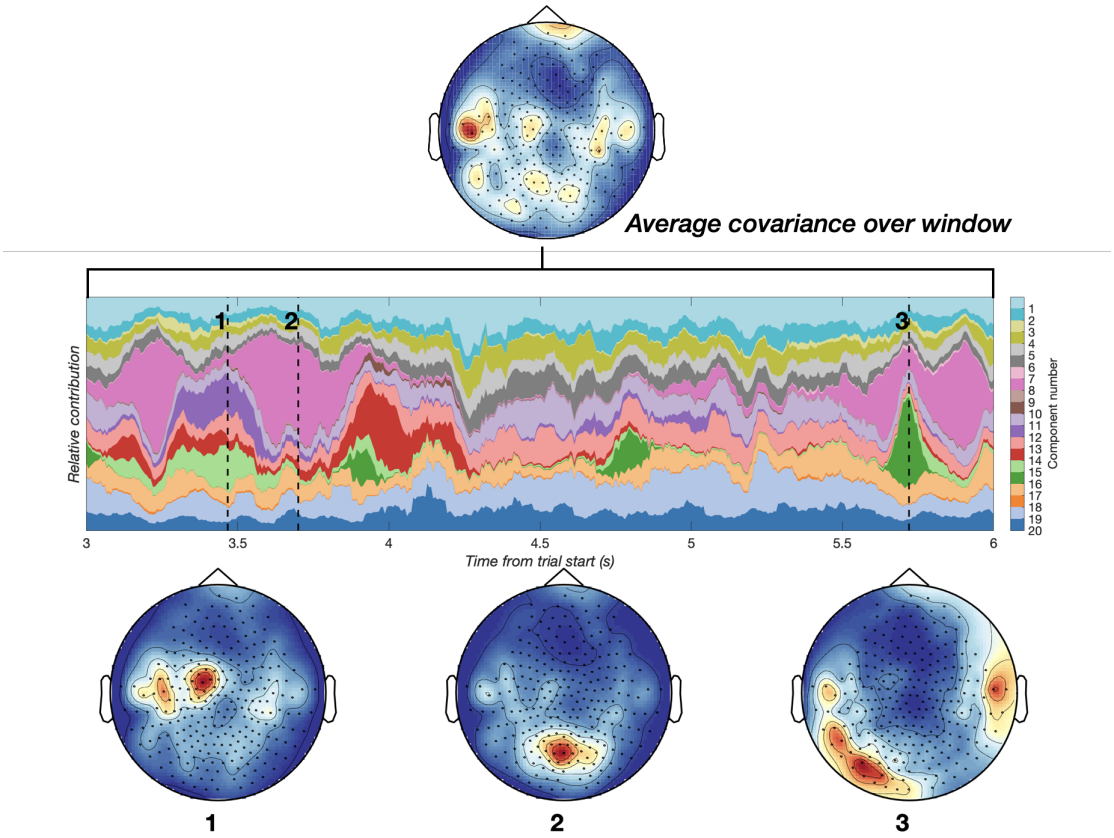


Figure 1.2: A potential confound in source localisation pipelines. Oftentimes it is assumed that statistical properties of MEG data do not change in time, and one averaged value is used in source inversion analyses. This idea is shown for a short three second segment on real MEG data; the time averaged topography (top) is one way of summarising the data, yet topographies are more refined when considered on a time point by time point basis (bottom). This can be decomposed as a mixture of components (colours).

lives. This includes keeping us safe from unwanted malicious emails, satellite navigation, voice recognition on our smartphones and chatbots designed to help us navigate customer services.

Whilst convolutional neural networks (CNNs) have seen a large uptake in image segmentation (Kamnitsas et al., 2017; Bao and Chung, 2018; Chen et al., 2018) and classification (Sahiner et al., 1996; Kim, 2014) tasks, recurrent neural networks (RNNs) naturally lend themselves more to time-series data. RNNs have seen application in speech synthesis (Zen and Sak, 2015) and recognition (Mikolov et al., 2010), handwriting recognition (Graves and Schmidhuber, 2009) and image generation (Gregor et al., 2015). One particular type of RNN is comprised of Long Short Term Memory (LSTM) (Hochreiter and Schmidhuber, 1997) modules, which have the added capability of recalling information from many time points prior, and the capacity to use this information to influence the current output of the network. More recently, work by Vaswani et al. (2017) showed that these convolutions and recurrent units could be replaced by a Transformer network, which centres on attention for sequence data. Importantly, the attention framework used in transformers facilitates parallelisation, massively increasing the efficiency of the network over C/RNNs and improves upon state-of-the-art scores in language translation tasks.

Some of the modern machine learning methods listed above have started to find

their way into neuroimaging analysis. In structural and functional MRI studies, this includes the segmentation of white matter lesions (Rachmadi et al., 2018; Sundaresan et al., 2021); the rapid automatic processing of cortical and subcortical brain regions from MRI scans (Henschel et al., 2020) and functional connectivity analysis (Pervaiz et al., 2021) - see Khosla et al. (2019) for an extensive review. In the M/EEG community, deep ML has been used to classify different stages of sleep from EEG data (Chambon et al., 2018), whilst Chehab et al. (2021) recently introduced a deep recurrent encoder network specifically designed for MEG data analysis. Again, the interested reader is pointed to a systematic review by Roy et al. (2019) for an overview of ML applications to electrophysiological data.

Structure and Objectives of this Thesis

The primary goal of this thesis is to investigate whether or not pre-existing approaches to source imaging can be improved by incorporating more intelligent modelling of sensor level data. Specifically, we ask what can be gained (or lost) by adopting numerous, distinct, source reconstruction weights[2], over the traditional setting where just one weights matrix is used. We do this by using different reconstruction weights at different points in time in order to better reflect the observed data - rather than just taking the average over the duration of an experiment.

A peripheral target of this work is to develop a new machine learning based framework to enable this temporal modelling. We would like to design a model which does not introduce arbitrary or restrictive assumptions about the observed data, so that we can reconstruct it with the highest possible accuracy.

Chapter 2 introduces the concepts underpinning signal production in MEG, modern scanners and source reconstruction. A formal mathematical description of the inverse problem is presented, and pre-existing techniques are interrogated and reviewed.

In **Chapter 3**, we review previous attempts at modelling time-varying neuroimaging statistics before introducing the new model-based machine learning architecture which facilitates the results presented in this thesis.

Chapter 4 ties the previous two chapters together and details how we can use temporal modelling of non-invasive signals to potentially better elucidate functional brain activity in the source space. Here, we introduce the concept of Temporally Adaptive SourcE Reconstruction (TASER), and present simulations to motivate the technique.

Work presented in **Chapter 5** attempts to exploit the spatial resolution which MEG has to offer. We reproduce results from a previously published study and show that MEG is capable of spatially disambiguating pinky and index finger activity in brain space. Importantly, unlike historic work, this is done on the non-phase locked oscillations which spontaneously arise post-motion cessation. We then apply various temporal models to the data and show that the best spatial resolution is achieved when a dynamic approach to source reconstruction is employed.

Penultimately, in **Chapter 6** we step away from the idealised experimental setup and assess how TASER performs in comparison to static methods in a more conventional MEG study. In this more challenging scenario, we reconstruct data to deep cortical sources in the presence of potential coregistration error.

[2] A *weight* within the context of source localisation tells us how much each of the sensor level recordings contributes to the inversion estimate. A more thorough definition is provided in Chapter 2.

Chapter 7 concludes this thesis by restating the key results learnt throughout. We also discuss key modelling and analysis steps taken along the way, whilst listing outstanding future work to be carried out by the next doctoral student and future MEG source imagers.



2

Magnetoencephalography: from source to sensor, and back again

Richard Feynman once asked:

“If, in some cataclysm, all of scientific knowledge were to be destroyed, and only one sentence passed on to the next generation of creatures, what statement would contain the most information in the fewest words?"

If one were to twist this question, insert the suffix “neuro" and point it directly at a neuroscientist, then we might expect to hear a response along the lines of:

“The brain is made up of billions of specialist microscopic components which communicate with one another via electrical and chemical signals."

In this chapter, we move beyond this single sentence and explore the biology and physics underlying the workings of magnetoencephalography (MEG) in detail. Starting at the sub-cellular level, we briefly introduce the fundamental building blocks of the brain - the neuron - and examine how they go about communicating with one another. We then discuss the most important cell types for MEG, contextualised by an overview of the physics underpinning the production of the signals which we measure. In the second section of this chapter, we introduce the problem statement of this thesis. We begin with a mathematical formulation of the inverse problem before reviewing historic attempts at solving it.

Signal origins in MEG

Neuron Fundamentals

Like other cells, neurons have a membrane, soma and other key organelles. They comprise two very noteworthy additional structures which facilitate intercellular communication: axons and dendrites.

Each neuron has one pre-synaptic axon extending from its cell body, although these can branch into collaterals. Axons are responsible for carrying electrical signals called action potentials from the soma via the axon hillock to the axon terminal at the cell’s periphery.

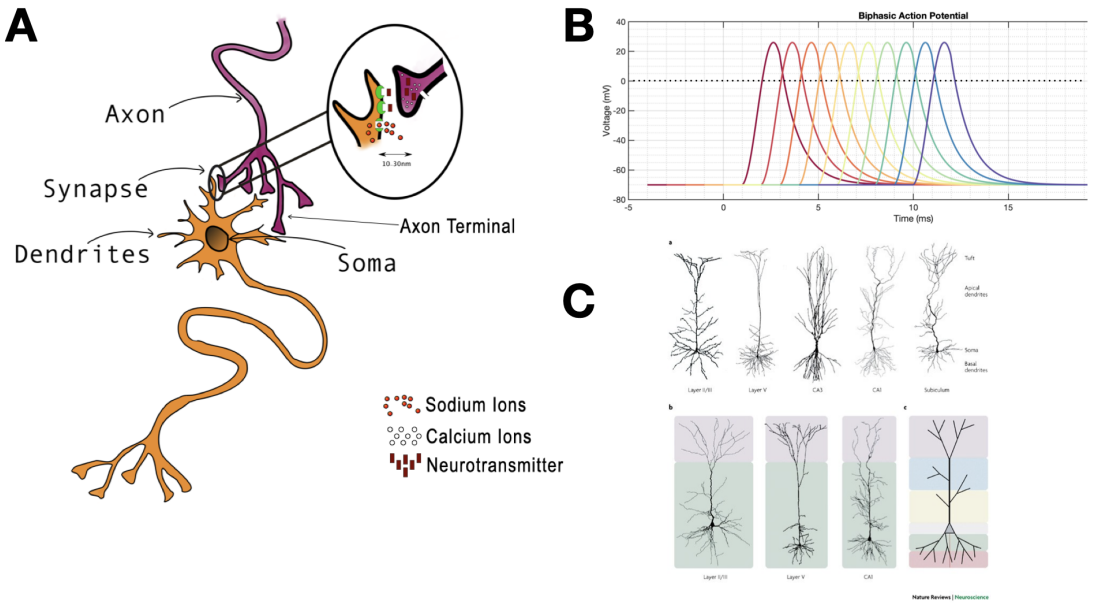


Figure 2.1: The basics of neuron structure and function. Panel A: a pre (purple) and post (orange) synaptic neuron. The key organelles are highlighted; the process of neurotransmission is shown as an inset. Panel B: the biphasic firing of an action potential over the course of approximately 20ms. The time lag shows how adjacent action potentials depicted in different colours may not necessarily lead to a net current flow. Panel C: the layering of human cortex, and its relationship to pyramidal neurons. Panels A and B were created by the author; Panel C is adapted from Spruston (2008).

Dendrites bifurcate into many smaller constituents away from the soma to drastically increase the cell's surface area and thus sensitivity to afferent signals from neighbouring cells. This makes up the post-synaptic part of a neuron and is where chemical signals are received.

Like the subatomic particle, most neurons do not actually touch, but are separated by a small (30-50nm) gap called a synapse. Intercellular communication is achieved when complex chemicals termed neurotransmitters diffuse from the pre to post-synaptic cell across the synaptic cleft. Neurotransmitters bind to ion channels on the dendrite (or the spines) of the recipient cell and cause them to open, resulting in a flood of ions entering the cell. These concepts are summarised in Figure 2.1A.

Action Potentials

Fortunately for the reader, your neurons do not tend to exist in vacuo. Instead, they lie in a bath of extracellular fluid which contains a swathe of organic material. When the neuron is at its resting potential, there exists a chemical concentration gradient across the encompassing 10nm thick membrane, a phospholipid bilayer. The gradient is established by the operation of transmembrane pumps fuelled by ATP (adenosine triphosphate). The result is a higher extracellular concentration of Na^+, Cl^- and Ca^{2+}, with more K^+ found inside the cell than out.

The laws of thermodynamics dictate that the sodium particles will diffuse towards the area of lower concentration (into the neuron). There are, however, other influencers at play here and one must also take into account the electrostatic forces arising from the charges on the particles. As there are more selective but non-gated potassium channels than there are sodium channels, the membrane is said to be more relatively permeable to potassium, and therefore it is these ions which are

more likely to traverse the cellular membrane as a consequence of the chemical concentration gradient.

This motion raises the relative electrical potential across the membrane. Eventually the exodus of positive ions leaves the intracellular space sufficiently negatively charged relative to the extracellular fluid, such that an electrochemical equilibrium is reached, and the potassium ions stop diffusing. Ex vivo experiments show that the voltage at this equilibrium point, the resting potential, is around $-60(\pm 20)mV$. Biophysical models derived by Nernst (1889) and later Goldman (1943) reach similar values.

The aforementioned ion channels lie on the surface of the membrane. These can be non-gated (i.e. always open) or opened/closed by electrical (“voltage gated") or chemical (“ligand gated") influence. If a sufficient number of positive charges arising from post-synaptic currents (see next section) reach the axon hillock and raise the membrane to a critical voltage (c. $-50mV$), gated sodium channels will begin to open, causing Na^+ to flow down the pre-existing concentration gradient into the axon. This influx of positive charges will begin to depolarise the membrane potential. The increase in potential results in the ion channels opening completely, and a further deluge of sodium rushes into the cell, leading to a voltage spike and a peak membrane potential of roughly $+40mV$ being reached in a few hundred microseconds. The presence of these positive charges causes a cascade of ion channel openings down the length of the axon, such that the voltage spike persists all the way to the terminal buttons of the cell. This is termed an *action potential*. See Figure 2.1 B).

Immediately after the sudden local influx of sodium, the ion channels shut and an electrostatically driven efflux of potassium across the membrane causes the potential to become negative again. Additional active transport of potassium continues even after the resting potential has been restored and the cell becomes hyperpolarised. This reduces the probability of a second action potential immediately firing (a “soft refractory period") and prevents the action potential from backpropagating up the axon to the soma.

Neurotransmitters and post-synaptic currents

The arrival of an action potential at the terminal buttons of a cell stimulates the opening of Ca^{2+} channels. The subsequent higher intercellular concentration of calcium triggers the release of complex chemicals termed *neurotransmitters*. These diffuse across the cleft to the post-synaptic neuron.

Protein receptors on adjacent cells bind with neurotransmitters. This triggers the opening of ligand gated ion channels and a flow of charged particles into the post-synaptic cell. This flow of charge is referred to as a *post-synaptic current*.

An example of an inhibitory neurotransmitter is GABA (gamma-Aminobutyric acid). When GABA binds to the post-synaptic cell, it facilitates the entrance of Cl^- into the post-synaptic neuron. This makes the post-synaptic neuron’s potential slightly more negative and thus harder to depolarise. In other words, it reduces the likelihood of the post-synaptic neuron firing.

In contrast, excitatory neurotransmitters like glutamate facilitate the entry of positive ions (chiefly Na^+ and Ca^{2+}) into post synaptic neurons such that an action potential is more likely to arise. However, it is the sum over potentially thousands of post-synaptic currents which ultimately determine whether an action potential will fire in the subsequent post-synaptic cell.

The human brain

Zooming back out, the central nervous system constitutes the brain and spinal cord. The telencephalon is made up of cerebral cortex and subcortical structures which include the limbic nuclei and basal ganglia. These are the primary targets for MEG imaging.

The outermost layer of cortex is made up of a very thin (1-5mm thick) sheet of highly folded grey matter. The resulting peaks and troughs, respectively termed gyri and sulci, give the brain a very high surface area to volume ratio, allowing a large number of folds to fit inside the relatively small human skull[1]. This tissue type is made predominantly of neuronal cell bodies.

Unsurprisingly a lot more can be learnt about the structure of the brain with the aid of a microscope in a process known as cytoarchitecture. Using Nissl staining, this is how Korbinian Brodmann categorised the brain into 52 structurally distinct areas in 1909 (Brodmann, 1909). When viewed under the microscope, one sees that 90% of the brain and all of cerebral cortex is made up of neocortex - areas where there are 6 distinct layers of different cell types - and allocortex, where there are only 3 or 4 layers. Despite this common lamination property, the thickness and cellular organisation of layers varies considerably across the brain.

Pyramidal neurons

There are myriad cell subtypes in the human forebrain. Of the population of cerebral neurons, about two thirds are pyramidal (Bekkers, 2011).

These multipolar neurons are named after their pyramid like soma and are found in layers II-VI of cortex, but are most prominent in layers II, III and V. All pyramidal neurons have stereotypically long apical dendrites which terminate at the molecular layer of cortex and are oriented parallel to one another. Shorter basal dendrites are oriented isotropically near and around the soma. Pyramidal cells are found with both small and large soma; particularly large cell bodies are found in layers IV and V of cortex and are termed Betz cells. These are particularly prevalent in motor cortex (see work by Rivara et al. (2003) for a modern analysis).

Cells found in layer II of cortex tend to be small, and only make nearby cortico-cortical connections. In layer III, the activity of cells is modulated by inputs from the brainstem. Things start to get a bit more interesting in layer IV, where thalamic axons are shown to synapse (Constantinople and Bruno, 2013). The axons from layer IV then themselves go on to synapse on neurons in layer V. Hence the one-step-removed influence of the thalamus on the cellular activity of pyramidal neurons in layer V (the primary contributors to the MEG signal) makes the role of the thalamus crucial in electrophysiological studies. Thalamic axons also terminate on stellate cells found in layer V of cortex which then themselves synapse on level V pyramidal neurons. These concepts are summarised in Figure 2.1C.

The biophysics of the recorded MEG signal

We have so far introduced two sources of current flow in the brain in the form of action potentials and post-synaptic currents. Which one of these signals contributes most to our extracranial recordings?

First, provided that the local axonal geometry is approximately straight, the aforementioned action potential can be modelled as two oppositely charged back-to-back current

[1] Given that the surface area of cortex is approximately $2000cm^2$, a back of the envelope calculation tells us that this would mean the *radius* of our heads would have to be $\sim 13cm$ in the absence of any cortical folding!

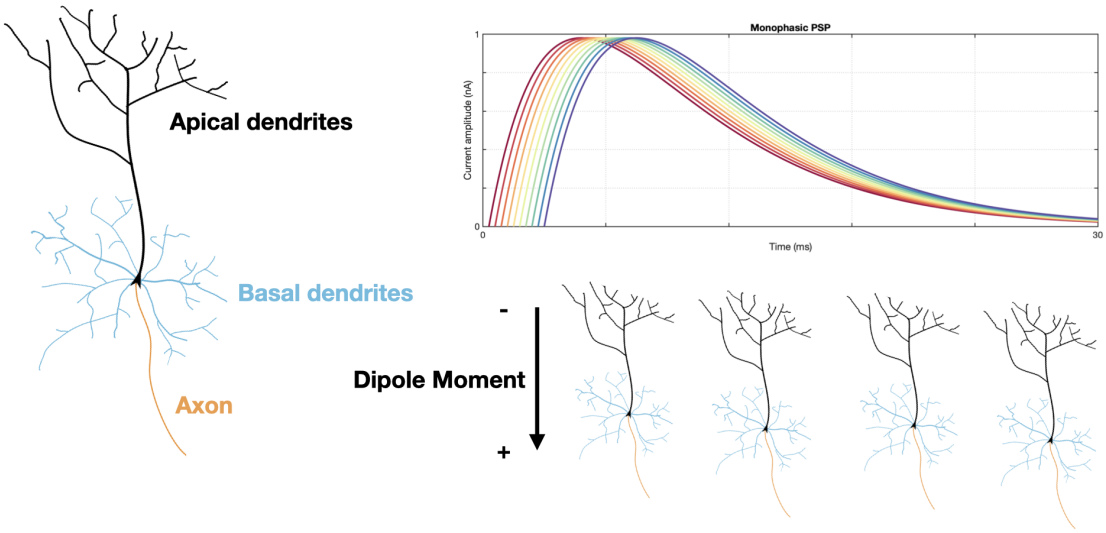


Figure 2.2: Signal generators in MEG. Left: schematic of a pyramidal neuron. Top right: monophasic firings of post synaptic potentials. Bottom right: the parallel alignment of pyramidal neurons leads to a net current flow and the superposition of magnetic fields.

dipoles (a “quadrupole") which propagates the length of the axon. Analytical expressions for such a current distribution tell us that the resulting magnetic field drops off with the cube off the distance from the source. In contrast, we model the excitatory/inhibitory post synaptic currents as a simple current dipole, pointing from source (apical dendrites) to sink (soma). The field arising from these currents only falls off with the square of the distance and is therefore more likely to contribute to the measured MEG signal.

The magnetic field associated with the firing of an individual neuron is too weak to measure with current MEG technology (Hämäläinen et al., 1993). We therefore require local firing synchrony in neuronal populations; the magnetic fields arising from $\sim 10^4$ neurons must superpose to give rise to a detectable signal ($\sim 10fT$) outside of the head. The profile of an action potential is biphasic, whilst invasive recordings tell us that post-synaptic currents are monophasic. We are therefore much more likely to encounter the constructive interference of small magnetic fields arising from the post-synaptic currents, where there is an inherent tolerance for imperfect phase alignment. Moreover, post-synaptic currents persist over 10s of milliseconds, whilst action potentials are much shorter lived. These concepts are shown in Figure 2.2.

The parallel alignment of apical dendrites found in pyramidal neurons is conducive to the constructive interference (superposition) of magnetic fields. Such configurations are not as prevalent in axons. Similarly, the symmetric currents found in basal dendrites are likely to lead to destructive electromagnetic field interference. The same argument can be employed to render the contribution of stellate neurons to the MEG signal unlikely.

To conclude this section, axoaxonal and axosomatic synapses are targeted by inhibitory GABA (Spruston, 2008), whilst axodendritic and axospinodendritic synapses are targeted by excitatory neurotransmitters from proximal presynaptic neurons. *It is for these reasons that the major source of the MEG signal is believed to originate from the synchronous current flow of thousands of excitatory post-synaptic pyramidal currents*.

Modern MEG systems

Now that we have described our sources of interest and how they are generated, we must turn our attention to recording them.

We start by considering two basic facts:

(a) The human brain creates magnetic fields, which are on the order of femtotesla.

(b) The modern world is an electromagnetically noisy place; the Earth's magnetic field is about a billion times larger than the fields that arrive at the surface of the head from the underlying brain activity.

This reality means that we are going to have to work very hard if we are to realise our ambition of successfully localising activity within the brain using MEG. Addressing these points in turn, to proceed we must:

1. Employ magnetic sensors which have the requisite sensitivity to detect the vanishingly small magnetic fields emanating from the brain.
2. Eliminate the dominating (and unwanted) signals present on the surface of the Earth which dwarf any neuromagnetic data.

Magnetic Sensors

There are only a handful of candidate devices which are sufficiently sensitive to detect neuromagnetic fields. These include super conducting quantum interference devices (SQUIDs) and optically pumped magnetometers (OPMs). Both exploit quantum phenomena in order to function.

SQUIDs comprise superconducting materials which show negligible electrical resistance at cryogenic ($< 11K$) temperatures. This is why SQUIDs are housed in a dewar of liquid ^{4}He. In MEG, these sensors are used in conjunction with large "pickup" and smaller "input" coils which are wired together in a circuit. As superconductors exhibit negligible resistance, and the circuit is a closed loop, there cannot be any total magnetic field. Therefore, any inbound magnetic field (flux) on the large coil will induce a second magnetic field in the smaller coil. The field through this smaller coil will have a larger flux and we detect this with a SQUID.

Superconducting or "BCS" theory (Bardeen et al., 1957) tells us that the induced magnetic flux incident through the superconducting loop is quantised. So long as a SQUID is kept within its linear operating range, it can be shown that the voltage across the SQUID is a periodic function of the number of magnetic flux quanta incident upon it. In other words, we have a proxy measurement of the flux incident on the pickup coil and we arrive at a magnetometer.

OPMs work by "optically pumping" a group 1 alkali metal vapour into a "dark state", such that any incident light on the vapour no longer interacts with the atoms. The gas is now in a spin polarised state; any magnetic field perpendicular to the optically pumping laser beam will induce a Larmor precession. Using a modulating field perpendicular to the path of the laser beam, integration of the Bloch Equations (Bloch, 1946) gives an expression for the polarisation of the gas in the presence of a non-zero magnetic field of interest. A polarimeter is then used to detect the polarisation of the probe light and the resulting voltage is related to the local magnetic field of interest (Tierney et al., 2019).

Crucially, the workings outlined above do not demand cryogenic cooling (in fact, Rb based OPMs are actually heated to $\sim 150°C$ to maintain spin polarisation in the so-called "SERF" regime). This makes OPMs an exciting up and coming technology, as they can be placed much closer to a subject's brain. However, the main technical

challenge that persists for OPMs is the requirement for there to be a (close to) zero background magnetic field. Just like the SQUID, this means that they need to be housed in a magnetically quiet environment to function.

Magnetic Shielding

In general, the most effective way of dealing with unwanted noise is to eliminate it prior to acquisition. The most important (and expensive) form of shielding in MEG is the magnetically shielded room (MSR). These are formed of interleaved layers of highly conductive and permeable materials. The former (typically copper or aluminium) are present to stop the effect of induced magnetic fields and the latter (a “mu" nickel-iron alloy) inhibits ferromagnetic fields.

Whilst imperfect (primarily due to cable guides, sharp discontinuities at the corners, and magnetic flux becoming trapped inside), MSRs are highly effective and capable of reducing the remnant background field to the order of nanotesla (Taulu et al., 2019). Such a reduction in noise is generally adequate for SQUID based systems, so long as the room itself is kept free from vibrations. OPM arrays stand to benefit from additional “active" shielding to further reduce any interfering field (Boto et al., 2018). One way of achieving this is by configuring current carrying coils to cancel out the remnant field due to the Earth, but also the spatial derivatives of any residual fields. This is what keeps an OPM within its linear operating range and facilitates small voluntary head movements, although alternative numerical methods are now being entertained (Seymour et al., 2021).

Even after we have reduced the noise floor of our system by several orders of magnitude, some noise invariably remains. Intelligent configuration of the aforementioned flux sensors can maximise the quality of our recorded data.

Sensor Configurations

Magnetometers pickup magnetic flux directly. As such, they are very susceptible to artefactual sources, but remain more sensitive to deeper brain areas. The other sensor type used in MEG, gradiometers, come in three flavours: planar, axial and synthetic. These detect the spatial gradient of magnetic fields in different (real and synthesised) planes. As such, assuming that most nuisance sources are distant from the brain, these sensors are less prone to interference, albeit at the expense of some depth sensitivity. Figure 2.3 shows the sensitivity profiles for an Elekta and CTF151 channel system.

Having established our sources of interest and the tools we can use to measure them, we may turn to the problem statement of this thesis.

The Inverse Problem: Definition

We describe the mapping from intracranial source currents, $\mathbf{X}$, to measured extracranial magnetic fields with a linear model:

$$\mathbf{Y} = \mathbf{H}\mathbf{X} + \mathbf{E}, \tag{2.1}$$

where $\mathbf{Y} \in \mathbb{R}^{N_c \times T}$ are the T MEG recordings measured across N_c channels; $\mathbf{H} \in \mathbb{R}^{N_c \times N_d}$ is the lead field, which, given a specified forward model encapsulating an individual subject’s anatomy, represents what magnetic fields would be measured at each sensor if there was a unit current source at each of the N_d dipole locations within the brain; $\mathbf{X} \in \mathbb{R}^{N_d \times T}$ and $\mathbf{E}$ represents noise. In words, the MEG signals are a linear combination of underlying signals of interest and external nuisance factors.

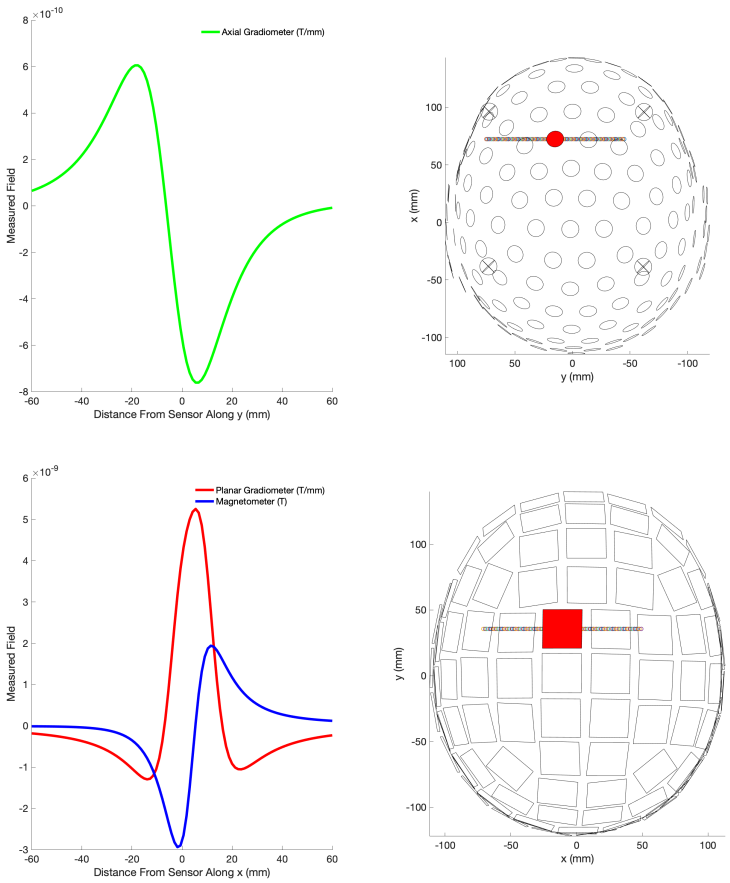


Figure 2.3: Investigating the profile of different MEG sensors for a 151 channel CTF system (green trace, top) and for a 306 channel Elekta system (bottom; red = planar gradiometer, blue = magnetometer). In both sensor arrays, an equivalent current dipole of fixed magnitude was placed pointing towards the front of the dewar, 5cm below a sensor of interest (highlighted in red). The location of the source was then systematically varied from left to right as shown by the small coloured circles (note the different scanner coordinate systems). The sensitivity of an axial gradiometer peaks adjacent to the source of interest but not above it. The same is true in the magnetometer case; only the planar gradiometer is most sensitive when directly above a source.

The goal of all source reconstruction algorithms is to find neuronal currents, $\hat{\mathbf{X}}(\mathbf{r}_i, \theta_i)$, with orientation θ_i at location $\mathbf{r}_i$. Solving Equation (2.1) is inherently ill-posed as no unique solution exists (Helmholtz, 1853). Furthermore, we typically attempt to reconstruct the time courses for thousands of sources with just a few hundred MEG sensors, rendering the system grossly under-determined. This also makes the lead field matrix highly rectangular, and so finding a numerically stable inverse is not guaranteed. These facts are all synonyms for the fact that the M/EEG inverse problem is insoluble.

Progress is made with the introduction of model constraints, for example smoothness (Hämäläinen and Ilmoniemi, 1994; Pascual-Marqui et al., 1994), sparsity (Matsuura and Okabe, 1995; Gorodnitsky et al., 1995a), or some combination of the two (Haufe et al., 2008).

Prerequisites in solving the inverse problem

Equation (2.1) tells us that to find an estimate for **X**, we also need an expression for **H**. This section details the ingredients which go into the formulation of this matrix.

Coregistration

In order for us to construct a lead field matrix and hence realise our ambition of source reconstructing our MEG data, we must know where a subject's head was with respect to the MEG sensors at the time of scanning. The quality of this coregistration can have a profound impact on the quality of any source localisation efforts as we shall discuss presently.

The typical goal of any multi-subject source space MEG study is to overlay a functionally derived statistical image on top of a "standard" anatomical image. Unless we are to resort to using one "template" head/brain/forward model for all subjects and thus compromise the accuracy of our modelling of the system, we are now required to turn to coregistration techniques.

Coordinate Systems

A MEG source space analysis pipeline has several different coordinate systems associated with it. We must equip ourselves with the relevant data to nimbly switch between these different spaces to achieve our aforementioned goal. The primary coordinate systems are as follows:

- "Native" MRI space. This is the space in which the anatomical MRI was acquired. Typically, the origin is determined by physiology by the radiographer at the time of scanning.
- "MNI" space. This is one of the most popular "standard" spaces upon which statistical images of the functional data are presented in publication.
- Head/scanner coordinates. This is also often referred to as "subject" space and is defined by anatomy and the MEG system which is in use.
- Polhemus/headshape space. The space in which digitised head points are acquired.

Moving between any of these coordinate systems is trivial by means of matrix multiplication(s). Our transformations are rigid body in nature (6 DOF) when moving between MEG and MRI coordinates[2] to maintain the veracity of the geometries (and hence physics) involved. Statistical images can then be non-linearly transformed to MNI space after any physics-dependent calculations are executed. The details of coregistration are typically taken care of by software packages for the regular MEG experimenter but are included in Appendix A: Coregistration for completeness.

Implications for source localisation

Coregistration is a very important and non-trivial step in any MEG source space study. Unfortunately, there is not much that we can do to correct poor quality data which are used in the coregistration process (i.e. the fiducial markers), and so it is crucial that good experimental protocols are followed at the time of data acquisition. To conclude this section, we take a look at the literature linking coregistration (and

[2] Some software packages such as SPM use non-linear transformations to warp a canonical/template brain to an individual subject's brain. This is only done once, and the transformed version of the template anatomy which is subsequently used in analyses is only ever subject to further linear transformations. Crucially, the non-linear operation does not change the number of vertices on the template mesh, which makes reporting results at the group level significantly easier.

thus forward modelling) errors directly to source inversion efforts.

First, it is informative to quantify the error one would expect to encounter in a typical MEG experiment. Hillebrand and Barnes (2011) approximate this figure to be between 5 and 10mm. Note that this is before we begin to consider any within session coregistration error due to head position (Stolk et al., 2013). Somewhat alarmingly, Liuzzi et al. (2017) showed within the context of a beamformer that Pearson correlations between a simulated and reconstructed timecourse dropped to between 0.7 and 0.5 within this expected error range. Efforts have been made to reduce the engineering challenge posed by coregistration with the uptake of foam headcasts (Troebinger et al., 2014; Meyer et al., 2017a; Barratt et al., 2018; Liuzzi et al., 2017), but this places further restrictions on the cohorts one might hope to image using cryogenic MEG, and has raised safety concerns in the past (Meyer et al., 2017a).

In the field of room temperature MEG, Zetter et al. (2018) showed that accurate knowledge ($< 4mm$ RMS error) of the location of sensors for an OPM MEG system is critical if the advantages of placing sensors on the scalp are to be realised. This is on a similar range of values reported by Hill et al. (2020). Interestingly, Hill et al. (2020) also showed that different experimental array setups resulted in substantially different errors in source modelling due to coregistration; within the context of a “helmet-like" OPM system where the coregistration errors are rigid body in nature, the correlation between simulated and reconstructed timeseries (again via beamforming) was markedly improved on a cryogenic system. Conversely, when an “EEG-like" flexible cap was used (where coregistration errors are free to be non-linear in nature), reconstruction accuracies dropped to scores comparable to a classic MEG setup.

The Forward Model

Now that the location of a subject’s brain with respect to the sensor array has been established, we can turn our attention to the biophysical forward model.

In order to solve the inverse problem, one must first solve the forward problem, that is deriving values analytically, or otherwise, for electromagnetic fields which would be recorded outside of the brain by a known neuronal source. A more complete treatment of forward modelling is provided in Appendix B: MEG and EEG Forward Models - we state the key take homes here.

The Single Sphere Head Model

The Sarvas formula for a homogeneous spherical conductor tells us the magnetic field, $\mathbf{B}(\mathbf{r})$, which we would expect to measure at a sensor located at position $\mathbf{r}$. It reads

$$\mathbf{B}(\mathbf{r}) = \frac{\mu_0}{4\pi\Lambda^2}\left[\Lambda\mathbf{Q}\times\mathbf{r}_Q - \{(\mathbf{Q}\times\mathbf{r}_0).\mathbf{r}\}\nabla\Lambda\right], \tag{2.2}$$

where $\mathbf{B}_0(\mathbf{r})$ is the magnetic field arising from the primary currents in an infinitely sized homogeneous conductor, $\mathbf{r}_0$ is the origin of the sphere, $\mathbf{Q}$ is an elemental current dipole at location $\mathbf{r}_Q$, μ_0 is the permittivity of free space and

$$\Lambda = |\mathbf{r}-\mathbf{r}_Q|\left(|\mathbf{r}-\mathbf{r}_Q|r + |\mathbf{r}|^2 - \mathbf{r}_0.\mathbf{r}\right). \tag{2.3}$$

Careful inspection of this Equation (and its derivation) tells us four key things about the single sphere model:

1. The solution is not dependent on conductivities inside the head.

2. Dipoles near the centre of the sphere do not contribute to the measured magnetic field.
3. Radial dipoles do not contribute to the externally measured magnetic field.
4. The solution does not depend on the radius of the sphere, only the position of the origin.

As helpful as Equation (2.2) is, the assumptions introduced along the way in its derivation limit its accuracy. Some of the shortcomings are obvious, whilst others are more nuanced and even demand numerical investigation.

The Overlapping Spheres Head Model

It has been known for a long time (Ilmoniemi, 1985; Hamalainen and Sarvas, 1989) that a single sphere fit to the inside of the skull can only be an accurate approximation for local areas of the head. For instance, the geometry of the skull makes the single sphere a good approximation for occipital regions, but decreasingly worse for temporal and frontal areas (Hamalainen and Sarvas, 1989). Errors were shown to be worse ($\sim 10-30\%$ in comparison to a more realistic “multilayer model") for deeper sources (across brain areas) where local skull profiles deviated the most from a sphere. Huang et al. (1999) addressed some of the shortcomings of the single sphere method and (at the time of publication) the computational overhead imposed by boundary element methods (BEM) by introducing the overlapping sphere model for MEG. Briefly, a local sphere is fitted to the local geometry of the skull within the vicinity of each sensor by minimising a least squares cost function. The Sarvas single sphere formula is then used to calculate the local lead fields after the radius and centre of each sphere has been found. This improved the accuracy of lead field approximations in the aforementioned troubling brain regions.

The Single Shell Head Model

The final forward model which we will consider here is the “single shell" method (Nolte, 2003). This uses the single sphere head model as a first order approximation, before adding spherical harmonics to improve the accuracy of the solution. This is achieved by imposing the boundary condition that the lead field is tangential to the surface of the volume conductor. Numerical work by Stenroos et al. (2014) showed that this approach produced lead fields very similar to more sophisticated BEM methods, at very little computational expense. It is for these reasons that we adopt the single shell head model in this thesis.

The choice of forward model complexity varies dramatically between laboratories across the field. As MEG investigators become more ambitious and begin attempting to reconstruct activity at brain regions with poorer spatial coverage, more elaborate modelling becomes important. This is certainly the case if the promise of studying the developing connectome is to be realised; the skulls of infants thicken with age, and it is known that BEM methods suffer from inaccuracies around thin regions (Stenroos et al., 2014) . Finally, it is worth highlighting the fact that researchers often rely upon inversion methods to assess the performance of their forward model. In other words, it is common for data to be simulated with various head models before they are reconstructed and used to calculate some accuracy metric. This can obfuscate the real performance of a forward model, with the results being more of a function of the imperfect inversion procedure than the forward modelling itself (Stenroos et al., 2014; Lalancette et al., 2011).

The Lead Field

After coregistration and a choice has been made as to which head model to adopt, the lead field matrix can be calculated.

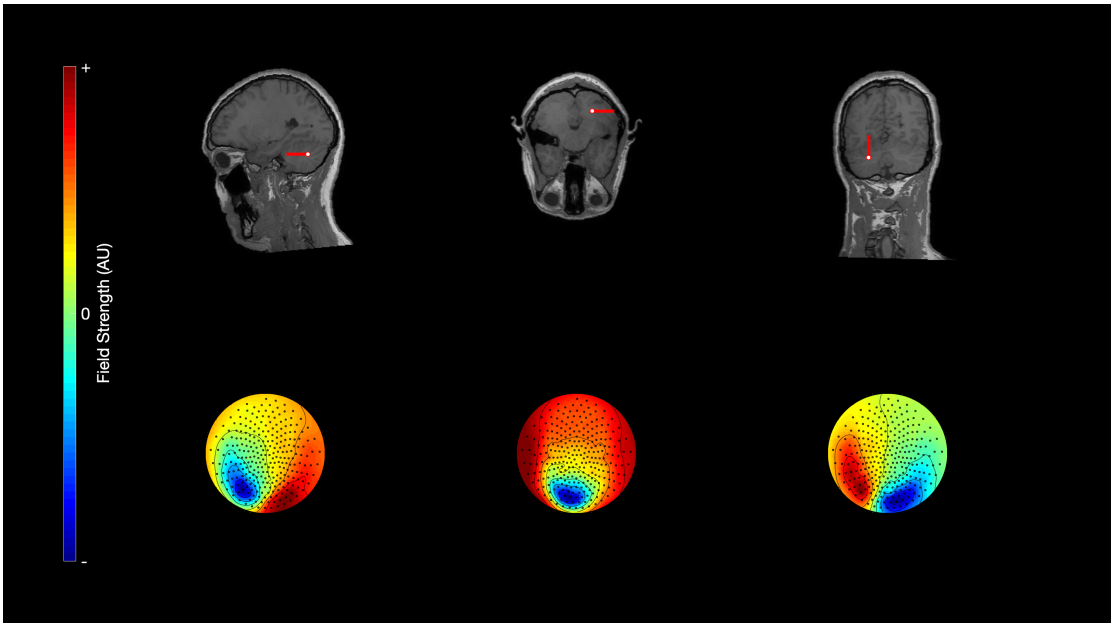


Figure 2.4: The effects of current orientation on lead field patterns. Top: equivalent current dipoles pointing in the positive x, y and z directions (in CTF head coordinates) give rise to dramatically different field patterns (bottom). Lead fields were generated for a CTF 275 channel system.

Formulism

Consider a current dipole of unit strength at location $\mathbf{r}'$ inside the brain. It is possible to write the contribution of that dipole to the externally measured magnetic field at each of the M sensors as

$$\mathbf{L}(\mathbf{r}') = \left[\mathbf{l}_x(\mathbf{r}'), \mathbf{l}_y(\mathbf{r})', \mathbf{l}_z(\mathbf{r}')\right], \tag{2.4}$$

where $\mathbf{L}(\mathbf{r}') \in \mathcal{R}^{M\times 3}$ and $\mathbf{l}_{x,y,z}$ represents the magnetic field which would be recorded along 3 orthogonal axes, given the unit source. More generally, we concern ourselves with N sources within the brain and so we write

$$\mathbf{F} = \left[\mathbf{L}(\mathbf{r}'_1), \mathbf{L}(\mathbf{r}'_2), \ldots, \mathbf{L}(\mathbf{r}'_N)\right], \tag{2.5}$$

where $\mathbf{F} \in \mathcal{R}^{M\times 3N}$. Formally, Equation (2.4) is the lead field matrix for a single current source and Equation (2.5) is the augmented lead field matrix for an entire source space. Note that here we have implicitly discretised the brain space so that we can use matrix notation.

Practically speaking, the inverse problem can at this stage be made “easier" to solve if we know (with good faith) the true orientation of each of the N sources. Once we have this information, we can swap our vector lead field, Equation (2.5), for a scalar lead field, $\mathbf{H} \in \mathcal{R}^{M\times N}$.

Source Orientation

The importance of source orientation in MEG cannot be understated; equivalent current dipoles pointing in three orthogonal directions give rise to profoundly different field patterns at the sensor level. This is demonstrated for a relatively deep source inside a CTF array of gradiometers in Figure 2.4.

The orientation of sources can be approximated either via good quality MRI scans *a priori* or by the functional data themselves. Both methods require an accurate

coregistration, although the former is more sensitive to potential errors (Hillebrand and Barnes, 2003).

As discussed earlier, histology tells us that pyramidal neurons are primary aligned perpendicular to the surface of the cortical mantle. When a high quality surface extraction is realised (e.g. via FreeSurfer), this knowledge can be baked into the forward model and used to constrain orientations to be perpendicular to the cortical surface. This is the default method employed in the SPM toolbox.

A halfway house approach between constraining sources to be perpendicular to the surface and solving the inverse problem in 3 orthogonal directions (i.e. Equation (2.5)) involves using a “loose" constraint on the dipole orientations (Lin et al., 2006a). This technique is implemented in the Brainstorm and MNE Python toolboxes and has the advantage that it reduces the sensitivity of the solution to the spacing of dipoles along cortex, which are a function of the local geometry. Anxieties surrounding coregistration error can also be relaxed.

Source orientation can also be “determined" by a heuristic[3]. These approaches are particularly popular within the beamforming community. The first method, synthetic aperture magnetometry (SAM) (Robinson, 1999), entails searching through the azimuthal angles between 0 and π for the optimal angle of each dipole which maximises the SNR at that location. This is only computationally feasible as beamformer weights at location $\mathbf{r}'_i$ are independent of all other source locations. As is the case with all beamformers, the maximisation criterion is over the SNR, as opposed to some goodness of fit metric; this metric is a function of the (sensor) data and noise covariance estimations, so inaccuracies in these terms can lead to erroneous orientations being derived. This also means that the orientations of sources at the same location may vary between sessions. Finally, it should be said that these learnt orientations are often not at all constrained by the underlying physiology, and so their values should not be overinterpreted. Despite these drawbacks, SAM has remained very popular since its introduction and is used routinely in paediatric epilepsy work (see Hall et al. (2018) for an excellent example).

An alternative data-driven optimisation approach entails using principal component analysis (PCA) to collapse three orthogonal current timeseries at one location to one “principal" timecourse which maximises the variance (Sekihara and Nagarajan, 2015). Fortuitously, the optimal angle of this synthetic dipole can be calculated efficiently via an eigenvalue decomposition of a linear combination of the lead field matrix and the empirical covariance.

Finally, one must consider what to do about source orientation when investigating subcortical structures (see Attal and Schwartz (2013) for a good discussion). Some brain structures are known to comprise highly organised cell columns, and so it is reasonable to impose orientation constraints when constructing the lead field. The hippocampus is one such exemplar region (Duvernoy, 2005), and has already been studied extensively with MEG (Meyer et al., 2017b; Tierney et al., 2021; Tzovara et al., 2019; Liu et al., 2019). Things are not always as straight forward for other areas: the amygdala comprises of pyramidal cells, yet these do not have any preferential direction (LeDoux, 2007). Finally, some regions of the brain such as the cerebellum are known to be highly structurally and functionally organised, yet remain hard to image (structurally) at the time of writing, precluding the construction of adequate mesh-based models (Samuelsson et al., 2020). In such cases, where any reliable information about orientation is absent, volumetric methods with three orthogonal lead fields at each dipole location is recommended.

[3]The quotation marks are employed here as the orientation of the current estimates in this method are not necessarily biophysical in nature, and are more of an abstraction.

Existing Inversion Algorithms

We now introduce and review popular inversion methods and discuss their merits. Broadly speaking, these can be split into three categories; equivalent current dipole approaches, spatial filters (including beamformers) and tomographic descriptions.

Equivalent Current Dipoles

Equivalent current dipole (ECD) methods assume that a small number of dipoles (circa one to five) can explain the observed M/EEG topographies for a single point in time. Estimations for the location of the source, $\mathbf{r}'$, and its angle, θ, have to be made, resulting in a non-convex cost function. Popular in the early days of MEG, applications of ECD methods are generally limited to event-related activity in very small time windows (e.g. Hari et al. (1980)), and cannot be used to image spontaneous rhythms across the brain. Nevertheless, in specific circumstances (e.g. finding epileptic sources (Rampp et al., 2019)), ECDs remain a popular and effective inversion scheme.

Beamformers

Beamformers are a data-driven spatial filter which opt to estimate a particular signal of interest arriving at some sensor array within the presence of noise (Van Veen and Buckley, 1988). A spatial filter exploits the fact that desired and interfering signals tend to originate from different positions in space, be it inside versus outside of the head, lateral versus medial temporal lobe, or different depths in an ocean (sonar, Albright Blomberg et al. (2013)). With accurate forward modelling, this spatial separation can be exploited to separate signals of interest from noise.

Let us momentarily assume that we have some *linear* way of going from sensor level signals to source space currents. We write our estimate for the source space current at location i as a combination of sensor level measurements:

$$\hat{\mathbf{x}}_i(t) = \mathbf{w}_{BF,i}^T \mathbf{y}(t), \tag{2.6}$$

where the spatial filter weights, $\mathbf{w}_{BF,i}$, vary across the brain. By sequential application of Equation (2.6), we are able to write down an estimate for the neural currents across the entire brain:

$$\hat{\mathbf{X}} = \mathbf{W}_{BF}^T \mathbf{Y}, \tag{2.7}$$

where $\mathbf{W}_{BF}^T \in \mathcal{R}^{N_d \times N_c}$ and $\mathbf{Y}$ is a matrix of the sensor level timecourse.

There is more than one type of beamformer. What differentiates one from the next is the constraint and assumptions which we place upon the solution. Typically, beamformer derivations begin by seeking to minimise the overall power of the source space solution, i.e.

$$\mathbf{w}_i = \min_{\mathbf{w}_i} \mathbf{w}_i^T \mathbf{C}_Y \mathbf{w}_i. \tag{2.8}$$

A solution is then found by the introduction of a constraint alongside a Lagrangian multiplier. One of the most popular beamformers is the linearly constrained minimum variance (LCMV) type, where a unit gain requirement is employed such that $\mathbf{w}_{BF,i}^T \mathbf{h_i} = 1$, where $\mathbf{h_i}$ is the lead field vector for dipole i. This ensures that a unit source at location i will be retained[4]. In such a scenario, the beamformer weights are found to be

$$\mathbf{w}_{BF,i} = \frac{\mathbf{C}_Y^{-1} \mathbf{h}_i}{\mathbf{h}_i^T \mathbf{C}_Y^{-1} \mathbf{h}_i}, \tag{2.9}$$

[4]This is in analogy to a finite impulse response (FIR) filter's weights being tuned to dampen any signals besides at the frequency of interest, where the signal remains at unity.

where $\mathbf{C}_Y = \mathbf{Y}\mathbf{Y}^T$ is the sensor level covariance. A derivation of this result can be found in Appendix C: Deriving Beamformer Weights. Mathematically, this is equivalent to maximising the SNR of the data.

It is not immediately apparent from the mathematical arguments presented here that the core assumption at the heart of the beamformer is that of statistical independence; the source space covariance is assumed to be a diagonal matrix and the correlation between any two dipoles is zero.

The concepts of a beamformer are summarised in Figure 2.6.

Historic Applications of Beamformers

The use of beamformers in the MEG community is incredibly extensive, to the point that the word is now almost synonymous with source reconstruction and it would be cumbersome to list all prior applications due to their sheer number. Beamformers have been used in all types of source space MEG studies including task data (Ishii et al., 2014; Hunt et al., 2019; Quinn et al., 2018; Perry and Singh, 2014), resting state paradigms (Brookes et al., 2011; Hipp et al., 2012; Colclough et al., 2016; Dimitriadis et al., 2018), functional connectivity investigations (Baker et al., 2014; Liuzzi et al., 2017; Boto et al., 2021), clinical interventions (Hall et al., 2018; Robinson et al., 2004; Stapleton-Kotloski et al., 2014; Ishii et al., 2008), the reconstruction of activity in deep regions (Cornwell et al., 2008; Tierney et al., 2021; Dalal et al., 2008; Groß et al., 2002; Lin et al., 2019; Quraan et al., 2011), spectral analyses (Seedat et al., 2020; Zich et al., 2020) and more.

Problems With Beamformers

Whilst very popular, beamformers by no means offer a perfect approach to source reconstruction, and present their own technical challenges. We discuss some of these now.

Covariance Estimation and Inversion

Equation (2.9) tells us that we need to estimate the sensor level covariance matrix from a finite window of data and then invert it.

When the noise levels in recordings are high, or when there is only a small amount of data available, covariance estimates can become poor. This often necessitates the introduction of regularisation to the covariance estimate to improve reconstruction accuracy (in time), but at the expense of spatial resolution in the source space images (Vrba and Robinson, 2000; Hillebrand and Barnes, 2005; Brookes et al., 2008). This can be achieved via Tikhonov regularisation (Tikhonov, 1943), PCA regularisation (Woolrich et al., 2011) or other means (Engemann and Gramfort, 2015).

Regularisation may also be required when the number of timepoints that we use to approximate the covariance is small. Brookes et al. (2008) showed that accurately estimating the covariance matrix is a function of the number of “useful" data points at our disposal, the SNR of the data, the bandwidth over which we estimate the covariance, and the number of sensors. The point to emphasise is that a sufficiently large time-frequency window is needed for good beamformer performance. This idea is particularly important and relevant to this thesis and we shall return to it throughout.

Finally, even after accurate estimation of the covariance matrix, we must still invert it. Mathematically, this can be highly problematic (impossible) when the matrix is

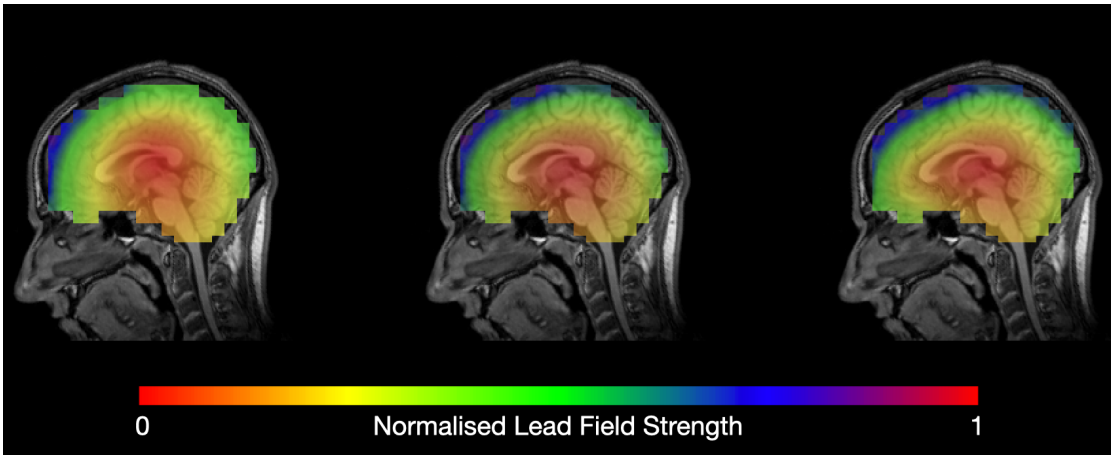


Figure 2.5: The impact of depth on lead field strength for different head models. Plots show the lead field strength (vector norm) for the single sphere (left), overlapping sphere (middle) and single shell head models. Notice the drop in sensitivity as we approach the centre of the head and how this is worst for the single sphere model. Also note that the more realistic models give greater sensitivity (blue colours) to dorsal and more deep areas. Values are normalised by their maximum value to ease visualisation.

rank deficient (due to T being smaller than $N_{channels}$ or as a result of preprocessing steps such as MaxFilter) or has a high condition number. Again, regularisation comes to our aid here. One option entails adding diagonal to the approximated covariance matrix to improve numerical stability.

Depth Bias

The unit gain constraint is integral to the LCMV beamforming philosophy; when the beam is “pointed" at source i, we want all that signal to persist through the filter. This can prove problematic when trying to reconstruct activity across the whole brain, as the norm of the lead field tends to zero as we approach the centre of the head. Consequently, the weights in Equation (2.9) become increasingly large in deep central brain areas. This is an artefact of the mathematics, as the norm of the denominator tends to zero and high variance is placed in brain regions which are poorly covered. Figure 2.5 shows this effect for each of the three head models discussed previously.

One way to circumvent this issue is to derive a different set of beamformer weights with another constraint. One popular choice is $\mathbf{w}_{BF,i}^{T}\mathbf{h_i} = ||\mathbf{h}_i||$. This is sometimes called the array gain minimum variance beamformer, as the beamformer gain is now equivalent to the inherent sensitivity of the MEG system and areas with poor spatial coverage (i.e. deep and/or radially oriented patches of cortex) are no longer artificially amplified by the unit gain constraint. Colloquially, we have a set of beamformer weights which are more representative of our inherent sensitivity to the brain. A formal derivation (Sekihara and Nagarajan, 2015) shows that the weights derived in this mode are in exactly the same form as Equation (2.9), except the lead field vectors are replaced with their normalised values, $\mathbf{h}_i \rightarrow \mathbf{h}_i/||\mathbf{h}_i||$.

Alternative depth weighting schemes involve dividing the beamformer weights by their norm (Luckhoo et al., 2014), an estimate of the sensor noise (Van Veen and Buckley, 1988), or involve constructing “pseudo" statistical images - see Vrba and Robinson (2002).

Coherent Sources

Theory tells us that beamformers cannot image coherent sources (Van Veen and Buckley, 1988; Brookes et al., 2004; Wipf and Nagarajan, 2009). Mechanistically, this issue arises from the fact that the lead fields from highly correlated sources superpose and the resulting sum from these vectors does not look like any of the true lead fields used in the construction of the beamformer weights. As a result, these recordings resemble noise (in the eyes of the spatial filter) and are thus attenuated.

In practice, this problem is found to manifest itself primarily in binaural auditory experiments where a bilateral evoked response is induced (Popescu et al., 2008), but has recently been identified as a problem in hippocampal beamformer studies, too (O'Neill et al., 2021). This problem, inherent to beamformers, is typically circumvented by summing the lead fields from brain regions which are hypothesised to give highly correlated responses (Brookes et al., 2007; Diwakar et al., 2011).

Tomographic Bayesian Approaches

The two previously discussed source reconstruction methods either attempt to fit observed data with an unrealistically small number of components, or drop any attempts to explain the data at all, and instead turn to maximising the SNR of voxel timeseries independently for each brain region.

Tomographic methods consider the influence of all source space dipoles *en masse* and attempt to explain the observed data, alongside some additional mathematical constraint. When formulated in a Bayesian framework, under the assumption of zero-mean normally distributed brain and noise data, the generalised maximum a posteriori probability (MAP) solution for the putative current sources is given by

$$\hat{\mathbf{X}} = \mathbf{C}_X \mathbf{H}^T [\mathbf{C}_n + \mathbf{H}\mathbf{C}_X \mathbf{H}^T]^{-1} \mathbf{Y}, \tag{2.10}$$

$$\hat{\mathbf{X}} = \mathbf{W}_{MAP} \mathbf{Y}, \tag{2.11}$$

where $\mathbf{C}_X \in \mathbb{R}^{N_d \times N_d}$ and $\mathbf{C}_n \in \mathbb{R}^{N_c \times N_c}$ are the source space data and noise covariance matrices, respectively, and $\mathbf{W}_{MAP}$ are the reconstruction weights (Lopez et al., 2014). See Appendix D: MAP Equation Derivation for a short derivation of this Equation. Inspecting Equation (2.10), one sees that the only user determined variables are the covariance matrices. These may either be specified *a priori* in the case of non-adaptive ("fixed prior") MAP estimates, or learnt/derived from the data in so-called adaptive MAP estimators. Regardless of what method is chosen, the veracity of these assumptions ultimately determines the accuracy of reconstructed currents and therefore must represent the reality of data.

Existing choices for $\mathbf{C}_X$ have ranged in sophistication from assuming that all voxels are i.i.d. (the minimum norm estimate (MNE) (Hämäläinen and Ilmoniemi, 1994)), to data-derived covariances as in the empirical based beamformer (EBB) (Belardinelli et al., 2012). When viewed in this guise, it is clear that the specification of $\mathbf{C}_X$ is what makes each reconstruction algorithm unique (see Figure 2.7).

The minimum norm solution (Hämäläinen and Ilmoniemi, 1994) to Equation (2.1) can be derived in a number of ways, be it via maximum likelihood, linear algebraic or formal Bayesian arguments (Hauk, 2004; Sekihara and Nagarajan, 2015). Whatever path we take, we find that $\mathbf{C}_X = \mathbf{I}$, which (in the absence of noise) is equivalent to using the generalised pseudoinverse to invert the lead field matrix (Penrose, 1955). To avoid over-fitting, estimates for $\mathbf{C}_n$ are used to regularise solutions. Hence MNEs represent a balance between goodness of fit to the data and a minimisation of total power. The implicit L_2 constraint in Equation (2.10) leads to smooth solutions.

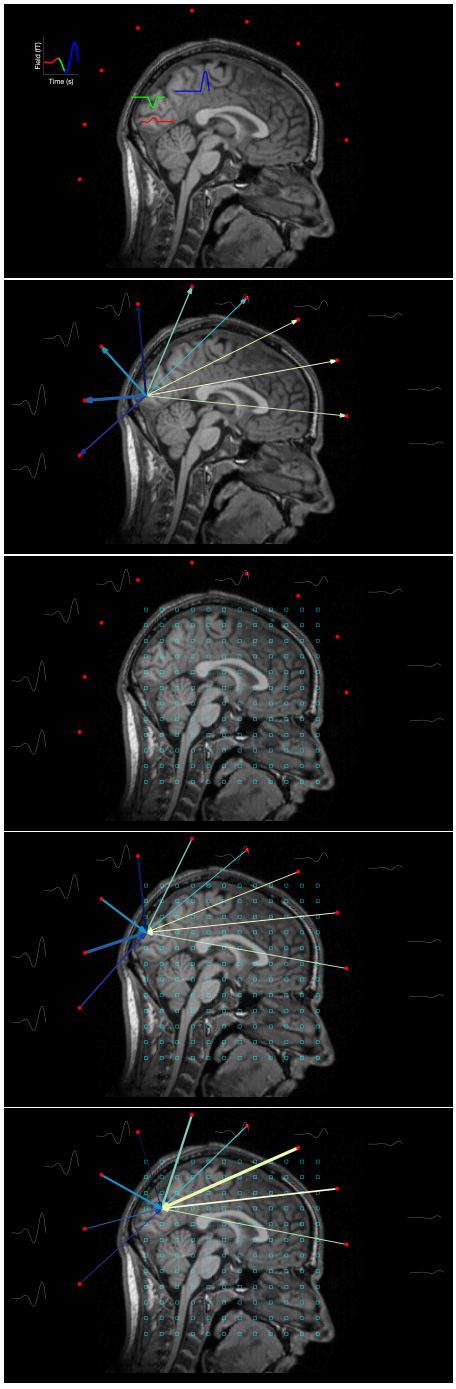


Figure 2.6: MEG signal production and source reconstruction. Row 1: neural signals at disparate brain areas produce magnetic fields which linearly combine to be measured as a superposition at the sensor level (red circles). Row 2: these sampled signals vary across space due to varying sensor/source separation and orientations. Row 3: when attempting to reconstruct brain activity, we limit our estimate to a finite number of putative sources. Typically this is done on an isotropic grid (cyan squares). Rows 4 and 5: source estimates are given at each brain region as a linear combination of the sensor observations. This step is carried out at each source location separately in the case of the spatial filter. Arrow widths and colours denote different weightings/magnitudes on an arbitrary scale.

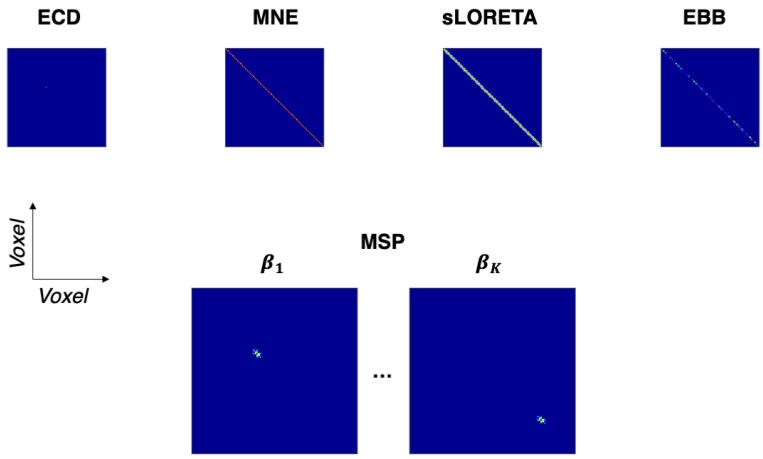


Figure 2.7: Distinction of source reconstruction algorithms can be understood as a different choice for the source space covariance matrix, $\mathbf{C}_X$. Five of the most commonly adopted structures are shown here. Abbreviations: equivalent current dipole (ECD); minimum norm estimate (MNE); standardized low resolution brain electromagnetic tomography (sLORETA); empirical Bayesian beamformer (EBB) and multiple sparse priors (MSP).

The i.i.d assumption does not marry up with the complexity of real neural activity. Moreover, MNE has a bias to superficial solutions (Lin et al., 2006b). This is because columns of $\mathbf{H}$ decrease for deeper voxels, and so more profound sources necessarily explain less variance. These issues were addressed by the weighted minimum norm algorithms, LORETA (Pascual-Marqui et al., 1994) and dSPM (Dale et al., 2000). Here, local coherence between neighbouring voxels is encoded by including non-zero off diagonal terms in $\mathbf{C}_X$ and depth-compensation is carried out by normalising the columns of $\mathbf{H}$.

Friston et al. (2008) generalised all the aforementioned approaches and used a sparse basis set of locally spatially coherent priors. Here, source covariance is parameterised by a weighted sum of typically hundreds of sparse priors:

$$\mathbf{C}_X = \sum_i^K \alpha_i \beta_i, \tag{2.12}$$

where $Q = \{\mathbf{I}, \beta_1, \ldots, \beta_K\}$ is the basis set, derived from a subject's anatomy in SPM's version of the algorithm. A parameter determines to what extent local patches of the cortex should be spatially coherent. An additional i.i.d noise prior is used to model the sensor-level noise. The goal of the MSP (or FOCUSS (Gorodnitsky et al., 1995b)) algorithm is to learn the set of α values which optimally explain the sensor space M/EEG data. This can be achieved by using Variational Bayesian (VB) inference, finding the MAP estimate for α via expectation maximisation, or by finding the MAP for the source currents directly.

The basis functions in Equation (2.12) are in high dimensional source space, $\beta_i \in \mathbb{R}^{N_d \times N_d}$. These are projected through a subject's lead field and the sensor space covariance is modelled directly,

$$\mathbf{C}_Y = \mathbf{H}\mathbf{C}_X\mathbf{H}^T = \mathbf{H}\left(\sum_i^K \alpha_i \beta_i\right)\mathbf{H}^T = \sum_i^K \alpha_i \tilde{\beta}_i, \tag{2.13}$$

where $\tilde{\beta}_i = \mathbf{H}\beta_i\mathbf{H}^T$. This massively reduces the parameter space to just $\tilde{\beta}_i \in \mathbb{R}^{N_c \times N_c}$. Note crucially that the α values in Equation (2.12) are the same as in Equation (2.13), such that learning these values in lower dimensional sensor space is all that is required to calculate Equation (2.10). Principal component analysis (PCA) or some other projection method is also used to reduce the dimensionality of $\mathbf{Y}$. Finally, the alpha values are given by $\alpha_i = e^{h_i}$, where h_i is a hyperparameter to ensure positivity in order to make the optimisation problem more tractable. Notable here is that unlike in the spatial filter regime, the model inference allows solutions to Equation (2.10) to self-regularise, and for regions of cortex which don't contribute to the overall variance to be turned off by the inference.

It is worth briefly listing some of the current limitations of the MSP algorithm found inside SPM. Most notable is the fact that MSPs is designed to work with short segments of data, and an implicit assumption is made that the covariance dynamics are stationary. This means that its application is currently limited to averaged evoked or induced responses, precluding the modelling of individual trial dynamics.

Reflection

So far in this chapter, we have introduced three separate source reconstruction philosophies, and the optimistic reader may now be keen to get to work creating source space estimates and answering some of neuroscience's deepest mysteries. A word of caution is advised here.

Let us plug our general linear solution to the inverse problem into Equation (2.1), temporarily neglecting the noise term:

$$\hat{\mathbf{X}} = \mathbf{WY} = \mathbf{WHX}. \tag{2.14}$$

Here we see that we do not expect to ever perfectly reconstruct our source space currents, unless somehow we end up with the result $\mathbf{WH} = \mathbf{I}$[5]. Further analysis (not considered in any detail here) shows that this a manifestation of the imperfect nature of the spatial filter/MAP estimates and is dubbed "spatial leakage". Briefly, activity at all other voxels in the brain contribute to the source estimate at voxel i, due to the columns of the lead field not being orthogonal[6].

This reinforces the fact that the M/EEG inverse problem is ill-posed and insoluble in the realistic settings that we use in modern MEG experiments. Nevertheless, there is ample literature detailing the historic successful application of modern source inversion techniques to study fast-paced neural dynamics across the brain. Finally, very recent work by Alberto et al. (2020) gives reasons for the MEG community to be optimistic about source imaging; using animal models they were able to validate the SAM beamformer approach by controllably activating deep (subcortical) structures via optogenetics, before successfully reconstructing activity to these areas.

Conclusion

MEG is unique in its ability to offer a direct window into human brain activity at sublime timescales in the form of spatial and frequency-specific images. In this chapter, we considered what the sources of encephalographic data are likely to be before showing that we are able to linearly reconstruct these signals back into brain

[5] Also note that the maximum rank of $\mathbf{WH}$ is $N_{channels}$, and so it is not generally invertible (Hauk, 2004).

[6] This may be an interesting area of research for OPM arrays in the future, where the lead field matrix can effectively be manipulated by intelligent sensor placement, thus theoretically allowing us to minimise the leakage into a pre-identified ROI. Whilst niche, this may be useful in some studies where a particular brain area is of interest.

space. We stressed the importance of the role that data covariance plays in source modelling, and showed how different ways of specifying this matrix give rise to different source algorithms. Whilst fundamentally ill-posed, solutions to the inverse problem can be found if one makes sensible assumptions about the underlying data. It is the “quality" of these user choices which ultimately determine how useful MEG source space estimates become.



3

A new approach to modelling

Introduction to chapter

Alongside the lead field matrix, the successful modelling of data and noise covariance matrices is paramount when tackling the M/EEG inverse problem. We propose that models that assume stationarity in the covariance are sub-optimal, and could lead to detrimental source reconstruction efforts. The novelty of this thesis centres around introducing dynamics into this modelling.

Here, we model time-varying data covariance as a linear weighted sum of basis functions, without making assumptions about which elements can contribute nor about the temporal dynamics, overcoming the mutual exclusivity and Markovian restrictions imposed by the alternative approach of the Hidden Markov Model (HMM). We use a unidirectional recurrent neural network (RNN) to model the temporal evolution of the weighted sum. The cyclical history provided by the RNN facilitates a richer context to the changes in covariance than a HMM. We use a bi-directional RNN to infer the model parameters. This amortises our inference and learns a mapping from raw sensor MEG recordings to dynamic descriptions of the covariance. This is achieved using stochastic gradient variational Bayes (SGVB) and the reparameterisation trick (RT).

Dynamics in the human brain

Our intuition tells us that cognition is a highly dynamic process. The fact that you can read this sentence in one quick instance, and then pick up your phone as you lose attention in the next, points towards our minds constituting dynamical brain systems. Neuroimagers, however, often make the approximation that the brain is static. This assumption can erroneously introduce spatial and/or temporal smoothness in results across various imaging topic (Vidaurre et al., 2019; Quinn et al., 2019; Seedat et al., 2020).

Analysis teams are moving away from this simplification of neural stationarity - it is now well established that the brain supports complex processes by recruiting brain-wide networks. These associations/disassociations have been observed to switch on the order of tens or hundreds of milliseconds in MEG (see Baker et al. (2014), for example).

We begin this chapter by reviewing some of the key literature, touching upon some

of the analysis techniques that we shall introduce in detail later. We will review efforts of dynamic covariance modelling specifically within the context of source localisation in Chapter 4.

Functional Magnetic Resonance Imaging Studies

As outlined in chapter 2, functional magnetic resonance imaging (fMRI) is fundamentally limited to imaging neural activity at slow (second) temporal scales, and as a consequence is blind to the nimbler activity continuously practiced by neurons. In spite of this, there is a large body of work evidencing non-stationary phenomena from a haemodynamical perspective. To list but a few key examples, Cabral et al. (2017) linked the switching dynamics of rsfMRI to cognitive performance in older age; Zhi et al. (2018) identified changes in dynamic functional connectivity (FC) profiles in individuals with major depressive disorder and Vidaurre et al. (2017) identified two sets of “metastates" which the brain was organised into. Most recently, Pervaiz et al. (2021) introduced the Multi-dynamic Adversarial Generator Encoder (MAGE) and stressed the difference between the dynamics of the mean brain activity and of the functional connectivity. Their work showed that by separating fluctuations in the mean activity levels from those in the FC, that functional connectivity is inherently more dynamic than previously thought, and is a better predictor of individual behavioural variability.

Electrophysiological Studies

Invasive and noninvasive electrophysiological methods do not suffer from the same temporal tardiness as fMRI (Hämäläinen et al., 1993). This makes them an ideal candidate for studying dynamic brain-wide phenomena at the fastest timescales. When using such imaging modalities, it is often beneficial to adopt analysis methods which do not assume that the underlying metric being studied is time-invariant.

For instance, De Pasquale et al. (2010) used sliding windows to better elucidate the dorsal attention and default mode resting state networks (RSNs) over a static counterpart. O’Neill et al. (2017) also used sliding windows to analyse motor and working memory task data, showing that brain network connectivity is modulated by action on a sub-second timescale.

Baker et al. (2014) championed and introduced the use of a Hidden Markov Model for MEG analyses. Their seminal work showed that dynamics in brain activity can be (statistically) represented as a sequence of several different transient “states", which are visited on the order of 100-200ms. This was one of the earliest studies to reliably tap into the fastest timescales underpinning whole brain organisation. Crucially, the spatial maps associated with these states resembled RSNs, which had been previously studied at slower time-scales, particularly in fMRI (Beckmann et al., 2005; Cole et al., 2014; Smith et al., 2009; Stevens, 2016).

As we shall see in the Hidden Markov Models section, the HMM is a very flexible overall analysis technique and has proven to be a popular tool in experimental neuroscience. This is because one can construct observation models in many ways and have many outcome measures of data dynamics (e.g. state lifetimes and intervals). Furthermore, identified states from task data have shown to be meaningful and informative (Borst and Anderson, 2015; Vidaurre et al., 2016; Quinn et al., 2018).

Existing approaches to modelling (dynamic) covariance

The above literature has highlighted the importance and prominence of dynamic modelling of human brain activity. We now detail some of the most common approaches

employed to achieve this.

A quick word on notation

When describing graphical models, it is common to use the symbols X and Z to denote observations and latent variables, respectively. In this thesis, we obtain source space estimates by modelling the sensor level data, $\mathbf{Y}$ .This is the symbol that we use to denote our observations (including any additive noise, see Equation (2.1) in Chapter 2). Z is commonly used when describing a Hidden Markov Model (HMM) state. We keep this notation, and attempt to draw the distinction between this and the new model introduced in this chapter by using the symbol θ to denote our time-varying latent variables. Also note that there is a tendency in the machine learning community to omit bold notation when talking about vectors or matrices. It should be clear from the context here whether or not a symbol represents a vector or matrix.

Sliding window analyses (SWAs)

Sliding windows have historically been one of the most popular ways of deriving time-varying statistics. Formally, sliding windows have a fixed width, w, which makes steps (forwards or backwards) of size δt. Often this step is small (e.g. 1 time point) such that the windows of data overlap; data within that window are then used to estimate some statistic e.g. the instantaneous covariance, $\mathbf{C}_Y$:

$$\mathbf{C}_Y = \frac{1}{W-1}\mathbf{Y}_{t:t+w}\mathbf{Y}_{t:t+w}^T \tag{3.1}$$

where $\mathbf{Y}_{t:t+w}^T \in \mathcal{R}^{N_{channels} \times w}$ is the (mean centred) data matrix and T denotes the transpose operation. The window is then slid along the timeseries and local estimates for the statistic of interest can be made. Once the N covariance matrices have been estimated for a session, some form of clustering (e.g. K-means (O'Neill et al., 2015)) can be employed to derive a state-like representation of the data, where modes of covariance can be seen to switch on and off periodically.

Drawbacks of the sliding window technique

The principal drawback of sliding windows revolves around the choice of hyperparameters. How long should we make our windows? Is our step size too big or too small? How should we summarise our statistics?

The other major hindrance is the choice between fine-grained dynamics and robustness. Smaller windows by definition give results with lower statistical power, and are more susceptible to outliers within a window of data (think of an eyeblink or heartbeat artefact in the MEG). Conversely, whilst longer windows are more robust to noise, their adoption introduces the risk of missing subtleties within the neural dynamics and, in extremis, we will end up with the static time-averaged metric.

Despite these drawbacks, sliding windows offer parity or even some advantages over more complex methods. First, they are easy to implement and are readily applicable across different data types. Second, when analysing data with inherently slow dynamics, there is often no reason to turn to more sophisticated approaches, as using longer windows does not lead to a loss in information. Third, and perhaps most importantly, windowing approaches can be adopted when deriving metrics that cannot be straightforwardly represented as a generative model in more formal

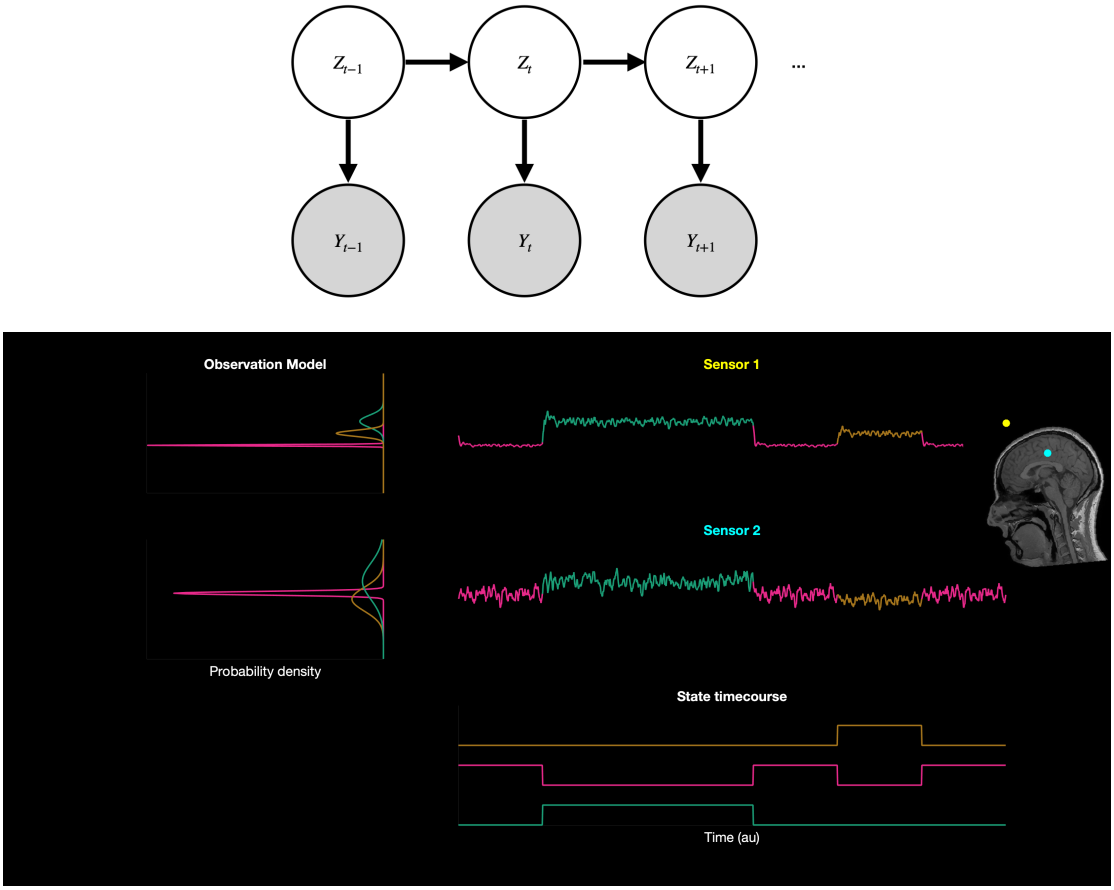


Figure 3.1: Schematic overview of an HMM. Top: graphical representation of an HMM. Observed data, **Y** (grey nodes) are generated by some latent process. The evolution of this latent variable follows a first-order Markov chain. Bottom: a 3 state HMM within the context of a (fictional) MEG dataset. The underlying generative state timecourse is linked to sensor level observations by an observation model with state-specific parameters, ϕ. Colours indicate which states were active. Adapted, with permission, from Quinn et al. (2018).

analyses.

Hidden Markov Models

SWAs are a heuristic approach to data analysis and lack a generative model, although remain popular. An alternative to studying neural dynamics is via adoption of a formal model. Hidden Markov Models (Rabiner and Juang, 1986) are one such option. HMMs are an example of a discretised state space model. The theory underpinning an HMM is the notion that some observed data, **Y**, can be adequately described by a (typically manageable) number of latent states, K. An HMM is able to explicitly model uncertainty within the hidden states by representing them as a sequence of mutually exclusive discrete random variables, Z, that evolve linearly according to a transition function $p(Z_t|Z_{t-1})$. This takes the form of a first-order Markov chain, with the state that is active at time $t+1$ only being dependent on the state which is “on" at time t.

This state description of the observed data is a convenient abstraction. We link the unobserved state sequence to the real observed data via an *observation model*, a probability distribution, conditional only upon the latent variable at that instance in

time, t. The values of these distributions are free to differ between states: in the case of a multivariate normal, each state can have a separate mean vector and covariance matrix. In this light, HMMs can be considered as a natural extension to simple mixture models - only now the mixture component for an observation has some (Markovian) temporal dependency (Bishop, 2006).

The modelling assumption, that the values of the latent state are conditionally independent of all states more than one timepoint prior, significantly reduces the computational complexity of the model. Hence, we may write:

$$p(Y, Z, \Theta) = p(Z_1; \pi)p(y_1|Z_1; \phi_y)\prod_t^T p(Z_t|Z_{t-1}; A)p(y_t|Z_t; \phi_y) \tag{3.2}$$

which is the joint probability distribution over the observed data, the hidden state time course and the parameters which define the HMM, $\Theta = \{\pi, A, \phi_y\}$. Respectively, these are the initial probability distributions, the $K \times K$ transition probability matrix (TPM) and the parameters pertaining to the values in the observation model.

Having written down the posterior distribution, we can start exploring the idea of inferring the terms above. Whilst our lives are made easier by the fact that the posterior can be written in this factorised form, we are still required to marginalise over the latent state dimension. To proceed, one now typically turns to variational methods to approximate the true posterior (Friston et al., 2007). Traditionally, this has come with the necessity of conjugacy between the variational and prior distributions to ensure tractability, thus limiting which observation models can be used. This limitation, along with the modelling assumptions of mutually exclusive state activation and the Markovian evolution of states, has not stopped HMMs from proving to be very powerful in the context of neuroscience and beyond (Kundu et al., 1989; Kupiec, 1992; Baldi et al., 2001).

Alternative HMMs

The state lifetimes modelled by a HMM are geometrically distributed. In other words, HMMs have a bias towards short state lifetimes. To mitigate this, one can turn towards Hidden Semi Markov Models (HSMMs) where the TPM is replaced by a distribution over the state lifetimes. This has led to longer state lifetimes being reported (in resting state EEG) than in Baker et al. (2014)'s work (Trujillo-Barreto et al., 2019).

Complementary to HSMMs, the Markovian constraint of a traditional HMM can be somewhat relaxed with the employment of a Higher Order Hidden Markov Model (HOHMM). The state at time t is now conditional upon the previous M time points, where M is the order of the HOHMM. This does come at the cost of increased complexity as for K states we end up having $K^{M-1}(K-1)$ parameters. In turn, this places us in a computationally prohibitive search space, as the computational burden scales with the square of the number of parameters Ghahramani (2001).

These desires, to have more temporal flexibility in the modelling of latent dynamics, alongside the restrictive categorical nature of the HMM, motivates us to find a new method for modelling non-stationary neural data.

Variational Recurrent Auto Decoder (VRAD)

Having briefly introduced some of the key techniques used historically to model covariance dynamics, we are now in a position to consider a new approach. We call

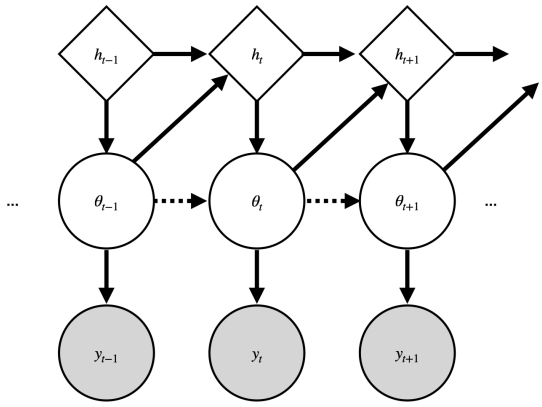


Figure 3.2: The generative process for the VRAD model. Sensor level observations are determined by the temporal evolution of the latent variables. Unlike in an HMM, the latent timeseries is governed not only by the previous value (dashed lines), but also by the hidden state of the RNN (diamond). Solid lines indicate explicit dependencies.

this a Variational Recurrent Auto Decoder (VRAD)[1].

Generative Model

The observed sensor level MEG data at time point t is governed by the generative model

$$\mathbf{y}_t \sim MVN(\mathbf{0}, \mathbf{C}_t), \tag{3.3}$$

where MVN is a multivariate normal distribution with mean vector, $\mathbf{0} \in \mathcal{R}^{N_c}$ and $\mathbf{C}_t \in \mathcal{R}^{N_c \times N_c}$ is the instantaneous covariance at time point t. We describe the data's time-varying covariance as a linear weighted sum of basis functions:

$$\mathbf{C}_t = \sum_j^J \alpha_{j,t} \mathbf{B}_j, \tag{3.4}$$

where $j = 1, 2, ..., J$. Note that the temporal evolution of the covariance is contained within the dynamic coefficients, α_t; the spatial basis functions do not vary in time.

The J dynamic coefficients at each time point t are associated with a set of (time-varying) latent variables, θ_t:

$$\alpha_\mathbf{t} = \zeta(\theta_t), \tag{3.5}$$

where $\zeta(x)$ represents some activation function. In this thesis we use the softmax function, described later in this Chapter.

The θ_t terms are parametrised deterministically by a model unidirectional recurrent neural network (RNN). For each of the J basis functions, their evolution is a function of the historic state time series, and also the hidden state of the RNN, $h_{<t}$:

$$p(\theta_j) = p(\theta_{j,0}) \prod_t p(\theta_{j,t} | \theta_{j,<t}) \tag{3.6}$$

$$p(\theta_{j,t} | \theta_{j,<t}) = N(\theta_{j,t} | \mu_{j,t}, \sigma^2_{j,t}) \tag{3.7}$$

$$\mu_{j,t} = g_{\mu\theta}(RNN(h_{<t}, \theta_{<t})) \tag{3.8}$$

[1] Marketing decisions carried out at the Oxford Centre for Human Brain Activity means that this approach may be called DyNeMo in the future.

$$\sigma_{j,t} = g_{\sigma\theta}(RNN(h_{<t}, \theta_{<t})) \tag{3.9}$$

where $g(.)$ are affine transforms and $\{\mu, \sigma\}$ are the parameters used for the distribution over θ.

Figure 3.2 shows a graphical representation of the generative model. As can be seen, unlike in a HMM, the RNN contains information about not only the previous latent variable values, θ_{t-1}, but also all other historical values, c.f. Figure 3.1. By extension, note that the temporal evolution of the covariance matrix no longer depends solely on the previous latent values, and that a plurality of basis functions being active at one point in time is now possible.

Nomenclature

We refer to the constituents of the VRAD model as “modes", “elements" and “basis functions" interchangeably throughout this work.

Inference

Our goal is to now learn the θ_t terms, and hence the time-varying contribution (quantified by α_t) of each of the covariance basis functions, $\mathbf{B}_j$. The flavour of source reconstruction philosophy/mode will determine whether or not we will either fix these basis functions *a priori*, or attempt to learn them (see Chapter 4).

One could opt to use maximum likelihood to predict these terms, yet accounting for uncertainty in our estimates has shown to lead to drastic improvements when modelling complex timeseries data e.g. Bayer and Osendorfer (2014). We therefore adopt a variational Bayesian approach to learn the latent dynamics. We are not fully Bayesian on the covariance basis matrices, nor on the weights and biases of the RNN, but instead rely upon point estimates made via gradient descent.

Given our probabilistic model, $p(Y, \theta)$, with observations, Y, we seek to infer θ. In the Bayesian parlance, this is achieved by calculation of the posterior distribution, $p(\theta|Y)$, which, in general, will not necessarily have a closed form solution. *Variational* Bayesian inference can help us here. We introduce an approximate posterior distribution, $q(\theta; \phi)$, to estimate the true posterior distribution, $p(\theta|Y)$. We now seek to minimise the discrepancy between these two distributions by learning the approximate posterior parameters, ϕ. Here we assume that the variational distributions are 1) factorised over time (the mean field approximation) and 2) normally distributed. Mathematically:

$$q(\theta; \phi) = \prod_t^T q(\theta_t|Y; \phi) \tag{3.10}$$

where

$$q(\theta_t) \sim N(m_t, s_t^2). \tag{3.11}$$

Now that we have defined our variational distributions, two outstanding issues remain.

Amortised Inference

Our first issue is that, for an EEG/MEG time series, there are still a drastically large number of parameters to learn in the variational distribution: a mean and variance term for each point in time and across each of the J basis functions. We therefore adopt an amortised inference approach. Briefly, we learn a single nonlinear mapping from the observed sensor level data, $\mathbf{y}_t$, to the (approximate

posterior) parameters which define the coefficient time series, $\{m, s\}$. We use a bidirectional RNN to define our mapping to the mean and standard deviation:

$$m_t = g_{m\theta}(RNN_{\longleftrightarrow}(h, \mathbf{y}_t)) \tag{3.12}$$

$$s_t = \rho(g_{s\theta}(RNN_{\longleftrightarrow}(h, \mathbf{y}_t))) \tag{3.13}$$

where $g(.)$ is another linear transform which is learnt (specific to the inference network, and not to be confused with the affine transform in the model network) and ρ is the softplus transformation, in place to ensure positivity. Note that we use $m_{\alpha_t} = \zeta(m_t)$ to calculate our dynamic weighted sum in Equation (3.4).

There is a degree of temporal causality in the model RNN, where historic values of θ_t influence the current estimate. In contrast, now that we are trying to infer these values, we are less concerned about any temporal causality between the observations and latent variational distributions. Using a bidirectional RNN, we are able to make the values for θ_t in the factorised posterior a function of both historic and future datapoints.

In this work we use Long Short Term Memory (LSTM) RNNs, which are described in detail later.

Cost Function

Definition

Returning to the discussion of the approximate posterior distribution which we introduced earlier, the log evidence of the observed data is given by

$$log(p(y)) = \int_\theta q(\theta;\phi) log(p(y)) d\theta \tag{3.14}$$

$$= \int_\theta q(\theta;\phi) log\left(\frac{p(y,\theta)}{p(\theta|y)}\right) d\theta \tag{3.15}$$

$$= \int_\theta q(\theta;\phi) log\left(\frac{p(y,\theta) q(\theta;\phi)}{p(\theta|y) q(\theta;\phi)}\right) d\theta \tag{3.16}$$

$$= KL\left(q(\theta;\phi)||p(\theta|y)\right) + \mathcal{L}(\phi;y) \tag{3.17}$$

where we used the product rule of probability on the second line; the chain rule on the third and the definition of the KL divergence on the fourth. This KL term quantifies how close the true and approximate posterior distribution match. The second term, $\mathcal{L}(\phi;y)$, is the variational free energy (VFE, but sometimes just denoted as F). As the log evidence term above is constant, we can see that maximising the fit between the true and approximate posterior distribution can be re-framed as maximising the VFE, up to a constant.

Expanding the logarithm inside Equation (3.17), the VFE is given by

$$\mathcal{L}(\phi;y) = KL\left(q(\theta;\phi)||p(\theta|y)\right) + \mathbb{E}_{q(\theta;\phi)}\left[log(p(y|\theta))\right]. \tag{3.18}$$

This is our analytical objective function, and we are now in the position to begin optimisation to find the best values of θ via gradient descent. The first term and its gradient in this Equation are trivial to calculate analytically when we use normal distributions for $q(\theta;\phi)$ and $p(\theta)$ as we do here. Obtaining gradient values for the second term can be tricky, as it involves an intractable integral:

$$\nabla_\phi \mathbb{E}_{q(\theta;\phi)}\left[log(p(y|\theta))\right] = \nabla_\phi \int_\theta q(\theta;\phi) log\left(p(y,\theta)\right) d\theta. \tag{3.19}$$

The Reparameterisation Trick (RT)

We must approximate the term above if we are to proceed. Upon calculation, this will allow us to calculate a gradient with respect to each of the network parameters and thus optimise values towards convergence via backpropagation.

We do this by sampling, a practice known in the machine learning community as the reparameterisation trick (RT) (Kingma et al., 2015). First, random values are drawn from a standard normal distribution: $\epsilon \sim N(\mathbf{0}, \mathbf{I})$. "Samples" from the current estimate of the posterior distribution are defined thus:

$$\theta^* = m_\theta + s_\theta \epsilon. \tag{3.20}$$

The expectation value above can therefore be approximated with this sample:

$$\nabla_\phi \mathbb{E}_{q(\theta;\phi)}[log(p(y|\theta))] \approx \frac{1}{S}\sum_{s=1}^{S} log(p(y|\theta^*)), \tag{3.21}$$

giving us an approximation for the log likelihood, i.e. the second term in Equation (3.18). Note that in our implementation of the inference, we typically use a sample size $S = 1$. The effect of sample sizes has been investigated elsewhere, although in the context of a very different application (Zhang et al., 2020).

Now that we have an analytical expression for the KL divergence and this approximation for the log likelihood, we can begin optimising our (modified) objective function,

$$\mathcal{L}(\phi; y) = KL(q(\theta;\phi)||p(\theta|y)) - \frac{1}{S}\sum_{s=1}^{S} log(p(y|\theta^*)). \tag{3.22}$$

We do this via stochastic gradient descent.

Figure 3.3 summarises the model and inference networks and their interaction.

Recurrent Neural Networks

RNNs are a special type of neural network which contain recurrent loops. This allows information from previous inputs in a network to persist as new input data are entered. Naturally this makes RNNs suitable for application to time-series data.

Figure 3.4 shows a unit of a basic RNN. It takes some input data, X_t, at a time-point t, performs some mathematical operations on it prescribed by A, and spits out an output, h_t. However, the looped arrow in the Figure means that information from the current input data, X_t, can interact with future input data, X_{t+1}. One sees how this interaction takes place more intuitively if the RNN is rolled out; results from the operations in A are fed into proceeding A units which act on future data inputs.

While basic RNNs do have some inherent memory, it is fundamentally limited to being short-termed. This is because such simple architectures as shown in Figure 3.4 do not have a means of explicitly highlighting important information from X_{t-N} inputs prior. Instead, basic RNNs will tend to be more heavily influenced by the most recent inputs.

Long Short Term Memory (LSTM) networks (Hochreiter and Schmidhuber, 1997) are a type of RNN that do not suffer from this short term memory problem, but are designed explicitly to be able to remember relevant information from many time points prior. This is achieved by using "input" and "forget" gates as well as a "cell state", which modulate the output and thus facilitates some "memory". One of the advantages of using formal model-based machine learning in a variational Bayesian

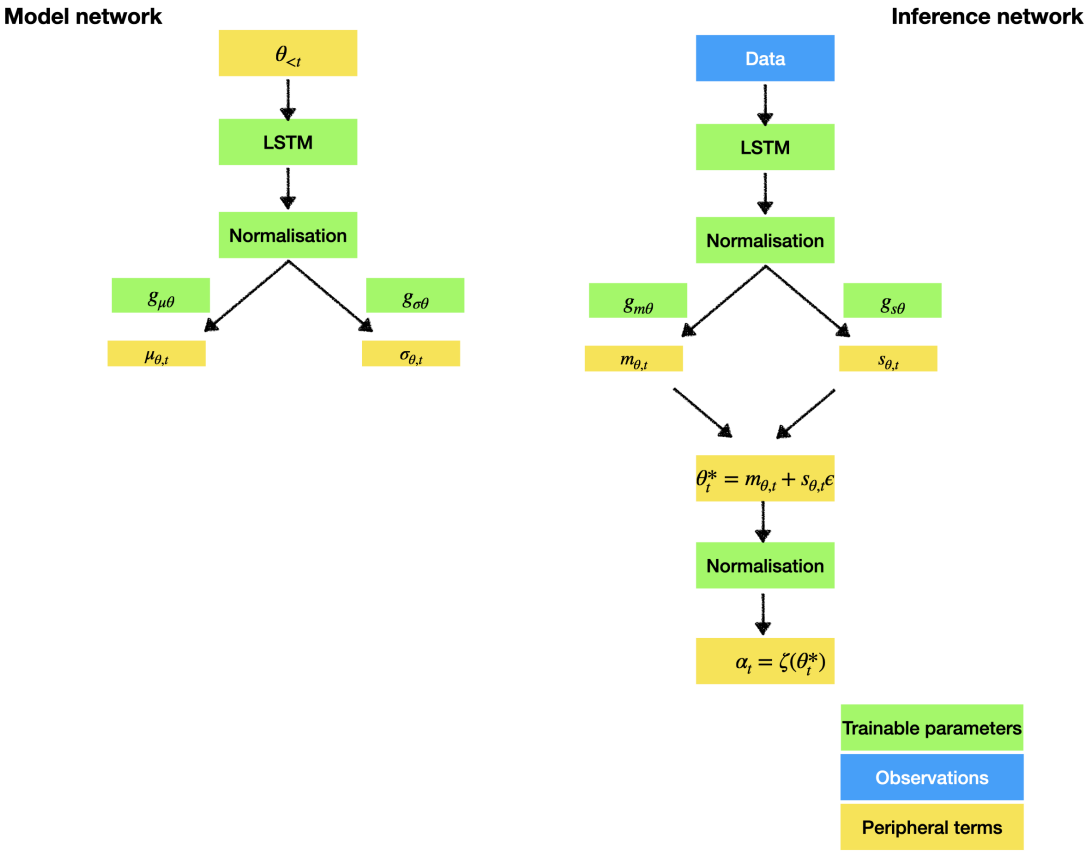


Figure 3.3: Model and inference networks shown in flowchart form. Trainable terms are shown in green, whilst auxiliary terms (parameterised by the learnt variables) are shown in yellow.

framework (like we do with VRAD) is that we can quite contently treat the RNNs used here as black boxes and be guided by a) our objective function and b) the final learnt parameters.

The final thing to say about LSTMs is that they do not have an infinite memory. Within the context of this work, this is to say that a latent variable at $t = 1$ will not have an influence on all future values. Instead, LSTMs have an effective memory or “receptive field", which is on the order of a few hundred timepoints (Geron, 2019). This equates to a few second’s of brain data when analysing MEG signals.

Implementation

The popularity of machine learning in recent years can give the false impression that training a model is as simple as downloading some pre-packaged code and hitting “run" (blog posts are particularly guilty for instilling this notion!). The reality, however, is that a lot of tips and tricks are required to successfully learn meaningful and non-trivial features from a dataset. In this section we detail some of the architecture and dark arts employed in this thesis, in an effort to reach convergence and avoid any numerical precision issues along the way.

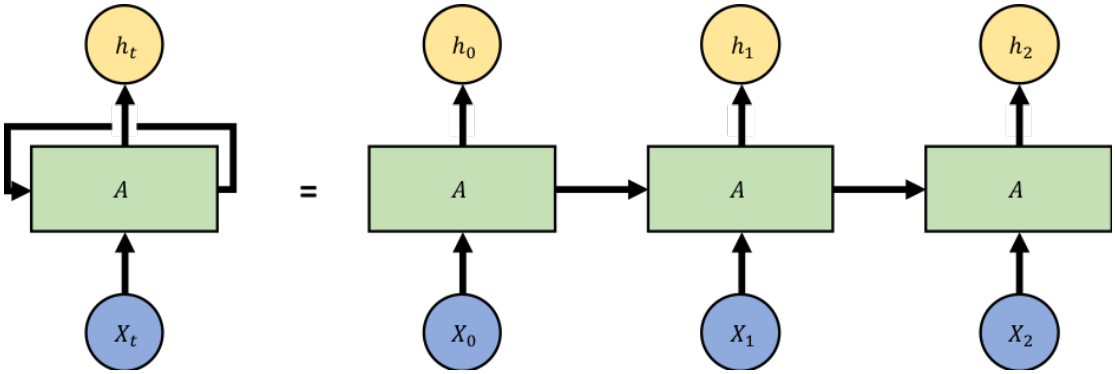


Figure 3.4: A basic recurrent neural network (RNN). Left: the basic concept of a RNN; data, X_t, are parsed into the network, where some mathematical operations within A transform the input into an output, h_t. Right: rolling out the network, the recursive nature of the architecture becomes more apparent, as output from one LSTM at a previous time point feeds into the proceeding one.

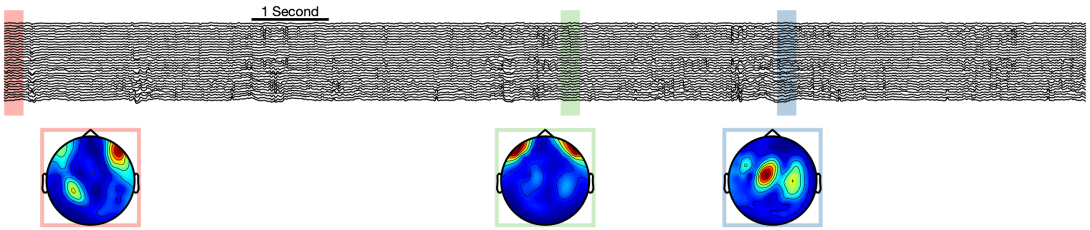


Figure 3.5: Initialising the basis covariance elements. Short segments of continuous data are chosen at random to populate the initial matrices. The variance for 3 randomly selected 250ms long segments are shown as topography maps.

Initialising The Inference

The weights and biases of both the RNNs and affine transformation layers were initialised with random samples from a Glorot uniform distribution and a zeros vector, respectively (the default settings in TensorFlow v2.5, used at the time of writing).

The optimal initialisation strategy for the covariance basis functions remains an open question for VRAD. We chose to z-transform our data (across the channels dimension, over all time) prior to feeding it into the network. A seemingly natural option is therefore to employ a set of identity matrices as the initial set of values for the basis functions. The intuition here is that, on average, this is the “correct" solution. However, as the goal of the framework is to tap into the covariance motifs at physiological (millisecond) timescales, there is a risk that such an approach risks trapping the network in a global minimum and/or enforcing a trivial one-state solution.

Empirically, we found that an effective approach was to take random (continuous) segments of the sensor level data and to calculate “noisy estimates" for the local covariance within these time periods. These were used to populate the initial estimates for the covariance basis functions at the start of inference. Tikhonov regularisation was employed to improve the numerical stability of these initial guesses. This can be thought of as using a SWA to initialise the inference. This is depicted in Figure 3.5.

Mini batch stochastic gradient descent

Once calculated, the gradient of the modified/approximate objective function can be used to update the parameters in the network (and the point estimates) towards the global minimum. Instead of using the entirety of the dataset to update our weights, biases et al., we opt to divide up the entire MEG timeseries of length T into sequences of length L. This can be used to make estimates for the gradient, $\nabla_\phi \mathcal{L}(\phi, \mathbf{y}_{t:t+L})$, and a subsequent update to the network parameters. This is called *stochastic gradient descent* (SGD) and is a popular method when data sets are large and/or occupy a lot of memory in RAM. Moreover, the stochasticity introduced by the sampling can help with moving out of local minima and aid the traversing of parameter space.

In mini-batch SGD, several samples are taken to estimate the loss before an update to the weights is made:

$$\frac{1}{b}\sum_{s=1}^{b} \nabla_\phi \mathcal{L}(\phi, y^s) \tag{3.23}$$

where y^s is 1 of b random sequences of the data. In this work we use the adam optimiser to update our parameters (Kingma and Ba, 2014).

We sought to pick values of $L \approx 200$ such that the benefits of the LSTMs could be maximised, i.e. such that the full history of the RNNs could be used (up to the receptive field).

KL Annealing

At the start of inference, the RNN parameters are pseudo-randomly initialised in both the model and inference networks. This means that the model network will initially produce quite naïve estimates for the α_t timecourse.

To sidestep this problem, we make use of an annealing phase at the start of training (Bowman et al., 2015), where a weight, η, is appended to the objective function:

$$\eta(n_{epoch}) = \frac{1}{2} tanh\left(\frac{\chi(n_{epoch} - n_{appeal})}{n_{appeal}}\right) + \frac{1}{2}, \tag{3.24}$$

where χ is the so-called “annealing sharpness". The loss function then reads $F = \eta KL - LL$.

KL annealing results in more emphasis being placed on learning the features of the data at the start of training (when effectively a maximum likelihood objective is used). However, as the dynamics of the data become more well known, the increasing annealing term acts to smooth out their profile, as the model network becomes more relevant. We then eventually end up at the classic VFE loss function.

Activation Functions

In this work we restrict the dynamic coefficients, α_t, to be strictly positive. We make use of a “softmax" activation function to accomplish this:

$$\alpha_t = \zeta(\theta) = \frac{e^{\beta\theta_i}}{\sum_j^K e^{\beta\theta_j}}. \tag{3.25}$$

This restricts the timeseries to not only be strictly positive, but also to sum to one at each instance in time. Our motivations for doing so are laid out in the discussion.

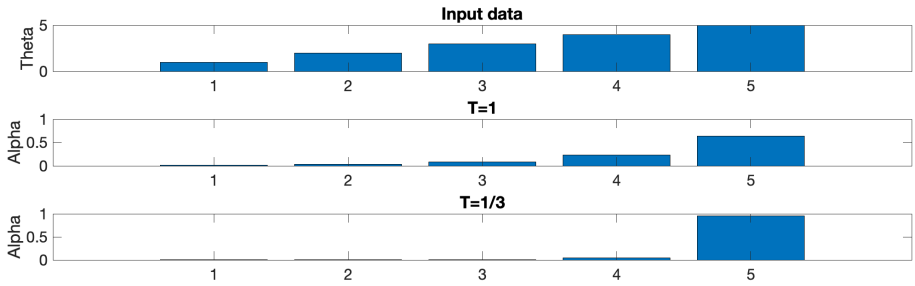


Figure 3.6: Behaviour of the softmax function for different temperatures. Top: input data, in our case the logits timecourse. Middle: outcome of the softmax transformation when $T = 1/\beta = 1$. Bottom: decreasing the temperature gives a “harder" or “more categorical" set of coefficients.

Temperature

We can define the “temperature" of the softmax function as $T = 1/\beta$ from Equation (3.25) above. Conventionally, the temperature is set to 1. However, altering the value of this parameter can fundamentally change the profile of the activation function, and therefore the (neuroscientific) interpretation of any results. Although a function of the input data themselves, lower temperatures typically result in a harder “more categorical" coefficient timecourse, with the output tending to the argmax of the input vector as $T \rightarrow 0$. In contrast, the coefficient time series becomes more “mixed" as temperatures increase. This effect is shown in Figure 3.6.

Arguably the most principled approach is to set the temperature to an initial value at the start of training and then let the data drive what this term should be. In this work we set the initial value of T to be 2.

Data Normalisation

Normalising data (via a z transform or otherwise) is an appropriate and often necessary appropriate step in modern machine learning (and in optimisation problems, more generally). This is for 4 key reasons:

- Gradients can become small/large enough to cause numerical instabilities if data are not appropriately scaled
- Convergence can be reached quicker when input data are normalised
- Data are considered more equally (i.e. features are all brought onto the same/a similar scale)
- Activation functions can become saturated if features are too large/small, c.f. Equation (3.25).

In this work, we employ layer normalisation, across each example/sequence (i.e. across features and not batches) to improve numerical stability (Ba et al., 2016). Inputs are z-scored and the network learns a scaling and shifting term to apply to the data (Geron, 2019).

Simulations

Showing the utility of a RNN-powered model

The result presented in this subsection was led by Chetan Gohil. All other work in this thesis was led by the author.

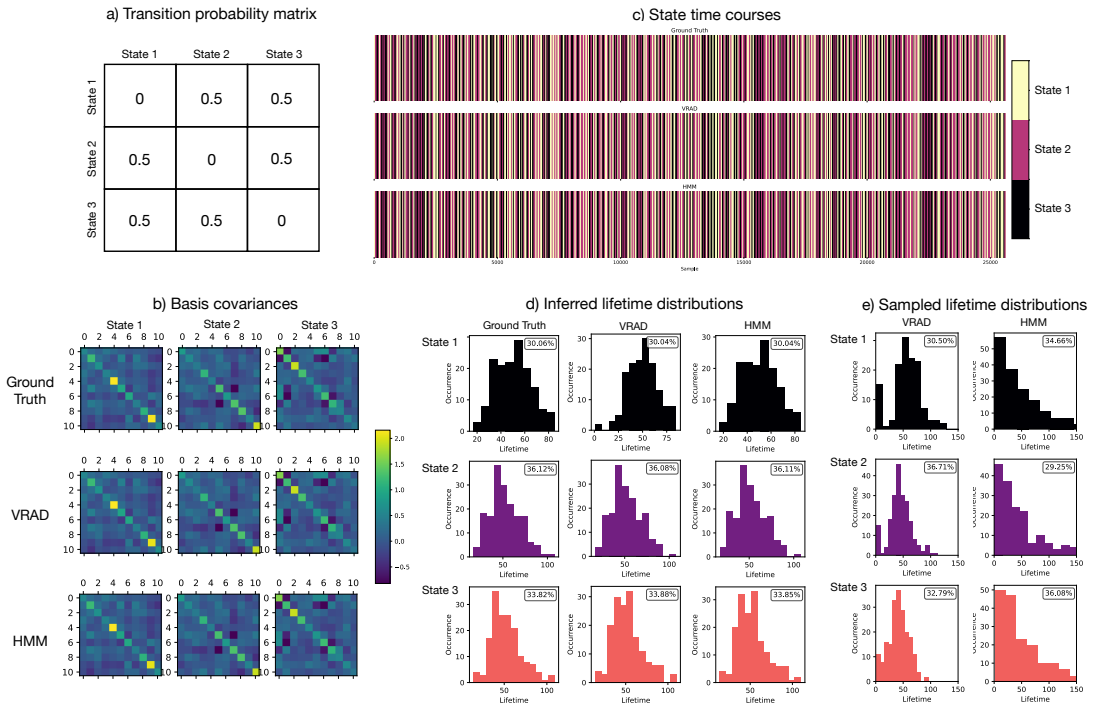


Figure 3.7: Demonstrating the increased modelling power afforded by a RNN over the HMM. Top: covariance matrices used to simulate data (arbitrary units). Middle: simulated (ground truth) and inferred state timecourses obtained from VRAD and the HMM. Notice how both methods obtain great results when the ground truth follows a binarised state timecourse. Bottom, left: histogram of the inferred state lifetimes for the two techniques plotted next to the ground truth for the $K = 3$ states (denoted by the different colours). Bottom, right: state lifetimes obtained from timecourses generated by the trained models. Notice the gamma distribution in the VRAD model (closer to the ground truth) and the geometric distribution in the HMM plots.

To demonstrate the increased modelling capabilities of VRAD over the HMM, we simulated data via an HSMM. If VRAD does indeed possess greater modelling power, then it should be able to capture the state lifetimes defined in the underlying ground truth model[2].

In this example, $25,600$ time points were simulated from a $K = 3$ state HSMM with a zero mean multivariate normal observation model for a synthetic 11 channel system. The state lifetimes were drawn at random from a Gamma distribution with a shape and scale of 10 and 5, respectively. The off-diagonal elements of the transition probability matrix (determining state transitions) were given the same value, making a transition to one of the other two states equally likely (see Figure 3.7a). The covariance matrices used to populate the data are shown in Figure 3.7b).

We fitted a $J = 3$ basis function VRAD model to the simulated data. Separately, we fitted a $K = 3$ state HMM with the default settings found in the OSL HMM toolbox (see https://github.com/OHBA-analysis/HMM-MAR). In the VRAD implementation, we use a sequence length of 200 data points when inferring the dynamic covariance. We used LSTM for the inference and model RNNs with 64 units and normalised the inputs to the RNN layer-wise; we did not normalise the θ layers in either network. The LSTMs in the inference network were bidirectional. We trained the model for 200 epochs with a learning rate of 0.01 and used the adam optimiser. KL annealing

[2]Recall that the lifetimes from an HMM will only ever be able to follow a geometric distribution.

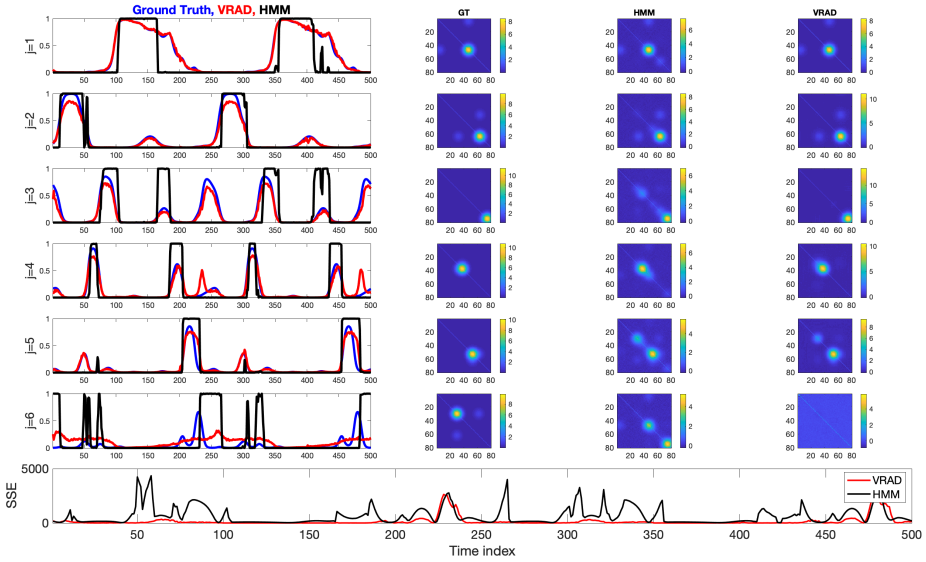


Figure 3.8: Demonstrating VRAD's ability to learn non-categorical latent descriptions of data. Blue traces in the left hand plot and first column of the covariance matrices show the ground truth used to construct the data. The first 500 timepoints of the dynamics are shown. VRAD recovers the latent descriptions of the data for the first 4 terms of the sum, before learning a mixture of elements 5 and 6 of the ground truth. This accounted for by the timeseries. In contrast, the HMM struggles to infer the matrix elements and is only able to approximate the dynamics with binary activations. This has a detrimental effect for the sum of squared errors (SSE) between the ground truth and inferred covariance matrices (shown in the bottom panel), where VRAD consistently outperforms the HMM.

was used according to Equation (3.24) for the first 100 epochs. In this simulation example, the temperature was deliberately held fixed at 0.25, to give a more categorical solution.

Figure 3.7c) shows the simulated ground truth and inferred state time course. Notice that both approaches do exceedingly well at learning the state dynamics (dice coefficients $= 0.9985, 0.9979$ for the HMM and VRAD, respectively). This is unsurprising, given the categorical ground truth.

Once trained, both the HMM and VRAD models can be used to synthesise data. A state time course can be sampled using the transition probability matrix in an HMM or the model RNN in VRAD. Figure 3.7e) shows the state lifetimes obtained when sampling from the trained models. Notice the gamma-like distribution in the VRAD column and its semblance to the ground truth lifetime distribution (first column). This highlights the new modelling capabilities of the RNN over a vanilla HMM.

VRAD can learn non-categorical descriptions of dynamic covariance

The other promised advantage of this new framework is the ability to model non-categorical contributions to data (co)variance.

To test this, we simulated a dataset with 51,200 timepoints for an 80 channel "MEG" system. The data were sampled from a zero mean multivariate normal distribution, with a covariance matrix which changed in time. The underlying ground truth covariance was constructed as a dynamic weighted sum of 6 basis functions. Crucially, these weights varied in an oscillatory pattern as can be seen in the blue

traces of the left hand column of Figure 3.8. The basis function elements which contributed to the total covariance were given by a Gaussian Kernel, multiplied by its transpose.

We kept the inference settings largely the same as the other example shown above, except for now learning the temperature during training (the initial value was set to 1); lowering the learning rate to $1e^{-3}$ and training for 500 epochs.

This particular simulation instance demonstrates how VRAD is better able to infer the latent dynamics as can be shown by the coefficient plots in Figure 3.8: VRAD (red) closely follows the ground truth timeseries for the first 4 states. Interestingly, here VRAD seems to “knock out" or “combine" states 5 and 6, as can be seen in the plots of the covariance (right hand columns) and compensates for this with the α_t timeseries. Notice how the HMM (black) is stuck learning more categorical-like descriptions of the latent dynamics.

When designing spatial filters, there is an argument to be made that one does not, nor should care about, how the time-varying covariance is learnt; so long as we get to the right solution, be it through a HMM, SWA or VRAD approach[3]. To test this, we plotted the sum of squared errors between the underlying ground truth $\mathbf{C}_t$ matrix and the reconstructed covariances derived from the HMM and VRAD as a function of time in the bottom of Figure 3.8. In this example, VRAD outperforms the HMM in this criteria, too.

Discussion

We have introduced a new generative model for neuroimaging timeseries data. Thanks to the stochastic inference techniques employed, we are able to make use of an RNN to provide more temporal context to the latent dynamics, and thus are able to demonstrate better temporal modelling than an HMM. RNNs apply (relatively) complex non-linear transformations to their input data. Employment of such a unit would not typically be possible in classic variational inference, due to constraints such as conjugacy. Finally, we used a multivariate normal distribution for the observation model in this work, but it is worth highlighting that more elaborate distributions could now be employed again without needing to worry about the classic variational conjugacy requirements.

We now explicitly model data as being generated by a linear weighted sum of covariance basis functions, meaning that non-categorical latent descriptions of data are technically feasible. This is particularly important for source reconstruction techniques such as temporally-adaptative multiple sparse priors (see Chapter 4), where feasibly hundreds of brain regions could be simultaneously contributing to the data covariance at any one instance in time; categorical models like the HMM are likely to just choose one of these basis functions as being active.

Whilst the non-categorical model can give more accurate results, it does come somewhat at the expense of interpretability. HMMs conveniently abstract data into mutually exclusive states, yet our data are now described as some complex combination of elements. This means that more effort must be put into the digestion of results, as demonstrated by the example in Figure 3.8.

The other major advance facilitated by VRAD is the use of amortised inference. Learning a parameterised mapping from the observed data to hidden variables means that VRAD could now be exposed to very large/long datasets without the explicit need to learn more terms. This is particularly useful for electrophysiological

[3]This indeed can be the case when modelling sensor level covariance for use in a beamformer, but is not when adopting a multiple sparse priors approach.

data, where we are blessed in the temporal domain.

A fair question to ask is “why did you choose that particular activation function?". Admittedly, our choice of a softmax, rather than a softplus or simple exponent (as adopted by Friston et al in the original multiple sparse priors paper (Friston et al., 2008)), comes about for somewhat arcane reasons. Anecdotally, we found that successfully training a softplus-powered network was harder. We hypothesise that this was due to the KL divergence term in the cost function “blowing up" due to the lack of a ceiling on the alpha terms. Finally, as the temperature in the softmax distribution goes to zero, the α_t timecourse becomes much more binarised. This can facilitate easier comparisons with other categorical models, such as the HMM.

We did not introduce any spatial priors to the network. As discussed in chapter 2, the success of beamformers is conditional upon good estimates for the empirical covariance. This is usually achieved by using many data points (i.e. by acquiring sufficiently long experimental recordings). This is sometimes difficult to achieve in practice due to, for example, clinical considerations (Cheyne et al., 2007). It may therefore be beneficial in future applications to make use of such informative priors (alongside a good initialisation) when a particular dataset is short and/or noisy[4]. Similarly, future work could look at being fully Bayesian on the parameters of the RNNs (Fortunato et al., 2017).

One important outstanding piece of work surrounds the run-to-run reliability of VRAD, particularly as Variational Bayesian methods are known to be sensitive to the way in which parameters are initialised. A second investigation as to what the optimal value of J should be is also required. Despite not being studied in detail here, other work (in prep) has shown that the negative VFE decreases monotonically with the number of basis functions, although the run-to-run variability appears to increase as J grows larger. Moreover, Skates (2019) showed in an analogous (formally categorical) model that VRAD-like approaches were able to reliably identify 8 or 9 unique states across subjects and different model initialisations. It should however be stressed that this was within the context of task MEG data that had already been source localised.

Conclusion

We have shown the increased capabilities of VRAD over an HMM and demonstrated that it can be used to learn a dynamic picture of covariance. Now equipped with this engine, and the theory underpinning source localisation methods from Chapter 2, we can turn our efforts to using such a tool for source reconstruction purposes.

[4]Note that whilst such an avenue begins to approach a MSP regime, “Bayesian beamformers" should not be confused as being a generative model source localisation approach (Friston et al., 2008; Sekihara and Nagarajan, 2015).



4

Temporally Adaptive SourcE Reconstruction (TASER)

Introduction to chapter

In this chapter we introduce the TASER philosophy to source analysis. We begin by highlighting two previous dynamic localisation efforts and discuss their key findings. We then show how the TASER approach can be thought of as a temporal extension to most of the popular established inversion algorithms.

Previous non-stationary source reconstruction attempts

Sliding Window Beamformers

Dalal et al. (2008) used sliding windows to construct time (and frequency) specific beamformer reconstruction weights, dubbing the technique “five dimensional neuroimaging". The researchers here put specific emphasis on the affects of low-frequency data dominating localisation weights alongside the confound of assuming stationarity in the covariance. This was demonstrated in simulation, where activity from a beta band source in motor cortex distorted images of weaker higher frequency sources when conventional SAM beamforming was used; adoption of multiple beamformer weights (for each window in time *and* each frequency bin) gave much more focal and accurate results and was able to disentangle the gamma and beta sources. This was only accomplished by using different covariance matrices for when each source (gamma and beta) were separately active.

The approach was further validated on real data, where results from a motor task across multiple subjects and invasive data from pre-surgical epilepsy patients were used to assess performance. For the healthy volunteer motor data, the non-stationary beamformer technique revealed “more focal" motor cortex activation as well as contralateral activity not seen in the standard beamformer approach. Interestingly, the dynamic beamformer also localised a clear gamma band source in primary motor cortex.

The authors highlight that this SWA is an embarrassingly parallelisable problem, which could be interpreted as a perk of the method. Nevertheless, the authors also admit that the analysis took ~ 20 minutes to compute on over 300 computer nodes, even for a relatively small number of short trials ($N \sim 200$, trial length $\sim 1500ms$).

Whilst access to cloud computing facilities are becoming more commonplace, this does place limitations on who can adopt the technique. It should be reiterated from Chapter 3 that as SWAs lack a formal model, they cannot be applied to unseen data. This means that the method employed here requires a new inference stage for each new dataset, placing another computational demand on the researcher. This is contrast to a VRAD-style approach, where a pretrained network can be exposed to new data at little extra computational demand (although we do not explore that possibility in this thesis).

The authors used relatively short ($100 - 300ms$) time windows to compute their dynamic estimates of covariance, potentially tapping into some of the quicker neural phenomena known to be observable in MEG. This gave rise to more focal source space images. However, the researchers still resorted to averaging the estimate for covariance over all trials. This may improve the reliability of the covariance estimate as the impact of any noisy trials will be averaged out, yet it has two key drawbacks. First, the covariance is still “static" over the width of the window. Second, and more importantly, by averaging over the trials dimension, any inter-trial variability exhibited by the brain will be averaged out. This has the affect of (partially) undoing any efforts to model the instantaneous covariance on a trial-by-trial basis. More advanced modelling techniques such as the HMM and VRAD do not suffer from this first confound.

HMM Beamformers

Woolrich et al. (2013) used a HMM to model sensor level covariance and derive a set of K distinct reconstruction weights. This work used a 13 state HMM. Their simulations showed that a dynamic approach was more robust in the presence of increasing measurement noise (in comparison to its static counterpart). Application to real data revealed bilateral activation of motor cortices following a seed based correlation analysis on resting state data, an effect not seen in the static beamformer. When applied to task motor data, the HMM beamformer elucidated constituent components of the motor network, whilst the static approach only found one large region of activity in either hemisphere.

This paper also contrasted the HMM with a SWA and showed that the latter approach gave very similar results to the static beamformer. Unlike the work by Dalal et al. (2008) however, the authors of this study used a very large window, on the order of a minute; when such large windows were used, any advantages in modelling the non-stationary covariance appeared to be lost. This reinforces the idea that one needs to model very fast (millisecond) changes in covariance if they are hoping to see an “improvement" in source reconstruction results.

As detailed extensively in Chapter 3, the categorical nature of HMMs and the resulting binarised state timecourses is restrictive, especially within the context of source reconstruction. For instance, it is likely that artefacts (e.g. the heartbeat) will coincide with genuine underlying brain activity. A HMM will struggle to model both simultaneously occurring events accurately. This is particularly problematic if we are to adopt a generative model based approach where dozens of brain areas may be simultaneously active. Realistically, this limits the adoption of a HMM as an inference engine to the dynamic spatial filter case.

The TASER Approach

As was shown in Chapter 2, most source estimates ultimately end up as a linear combination of the sensor level measurements. By introducing a temporal dependency on the data and noise covariance, the Equation used to reconstruct data for a given

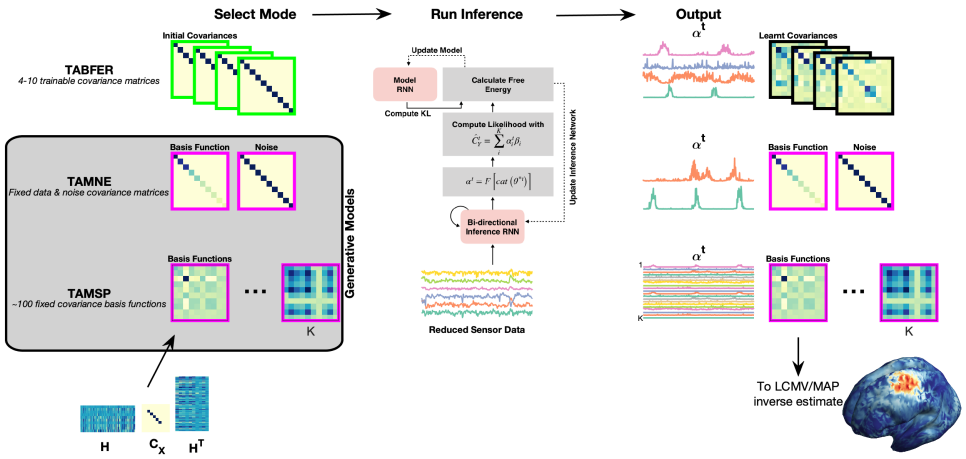


Figure 4.1: Schematic overview of the TASER algorithm. Left: covariances for the inference framework are prepared depending on the source reconstruction flavour of choice. For adaptive spatial filters, the covariance matrices are learnt (green squares). In generative models (pink squares) the K basis functions are fixed by the source space prior, $\mathbf{C}_X$, and the lead field, $\mathbf{H}$. Middle: reduced sensor data are fed into the inference RNN, which parameterises the α_t timeseries and thus the instantaneous covariance matrix for each point in time. Right: upon convergence, the approximation for the instantaneous covariance is used to calculate inversion weights for each point in time.

timepoint, t, now reads

$$\hat{\mathbf{x}}(t) = \mathbf{W}(t)\mathbf{y}(t), \tag{4.1}$$

where $\mathbf{W}(t)$ can be the LCMV beamformer weights (Equation (2.11)) or the MAP estimate (Equation (2.10)). We call this approach "Temporally Adaptive SourcE Reconstruction" or "TASER". The different flavours/setups of TASER are now discussed.

Temporally adaptive multiple sparse priors (TAMSP)

Here we extend the existing MSP algorithm and generalise it into the temporal domain, modelling the source data covariance as time-varying positive weighted linear sum of basis functions:

$$\mathbf{C}_X^t = \sum_i^K \alpha_i^t \beta_i. \tag{4.2}$$

We employ the same marginalisation step introduced in the original MSP work, and project the source priors to the sensor level with the subject's lead field, such that we model the time-varying sensor-level covariance directly:

$$\mathbf{C}_Y^t = \sum_i^K \alpha_i^t \tilde{\beta}_i \tag{4.3}$$

where $\tilde{\beta}_i = \mathbf{H}\beta_i\mathbf{H}^T$ are fixed basis functions. We also reduce the dimensionality of the data (in space) as in the MSP code, however we choose to use a PCA rather than a singular value decomposition[1].

[1] For ease of reading we do not change our notation after this step.

Temporally adaptive minimum norm estimate (TAMNE)

The minimum norm estimate can be considered a special case of Equation (2.10) where only one source space prior (equal to the identity matrix) is used. We therefore model the covariance as a simple two term sum:

$$\mathbf{C}_Y^t = \alpha_1^t \tilde{\beta}_1 + \alpha_2^t \mathbf{I}, \tag{4.4}$$

where the first term represents the data prior, projected to the source level, and the second term represents the noise regularisation (α_2^t is often given the symbol λ in the source recon literature, and isn't time-varying). Left with only two degrees of freedom, in such approaches an accurate estimate for the noise regularisation term is paramount (Dalal et al., 2008; Backus and Gilbert, 1970; Hincapié et al., 2016). Whilst some software packages prescribe values for the regularisation parameter, a Bayesian approach to the minimum norm solution can give an optimal value for this term, which changes on the properties of the data. This has been achieved using the EM algorithm (Sekihara and Nagarajan, 2015) or using restricted maximum likelihood (Lopez et al., 2014) in static implementations of MNE.

The rationale for being time-varying on the MNE is that we can expect the properties of our data to change in time. We should not therefore necessarily restrict ourselves to using just one value for α_2^t: ideally, noisy segments of data should receive more regularisation than cleaner ones.

Temporally adaptive beamformers (TABFER)

Finally, in the dynamic spatial filter variant of TASER, we adopt the same linear sum model, but this time infer the spatial elements of the weighted sum, as well as the dynamic coefficients:

$$\mathbf{C}_Y^t = \sum_i^K \alpha_i^t \mathbf{B}_i, \tag{4.5}$$

where the matrices $\mathbf{B}_i$ are trainable.

These concepts, for each of the different modes of operation, are summarised in Figure 4.1.

Dimensionality Reduction

Upon convergence, we will have learnt up to T unique $\mathbf{C}_Y^t$ matrices. Interpreting and digesting such a large number of results (in contrast with just K covariance matrices for a HMM) is challenging. Within the context of source reconstruction, the more practical issue of calculating up to T different sets of reconstruction weights presents itself. Each $\mathbf{W} \in \mathcal{R}^{N_d \times N_c}$ matrix amounts to a few megabytes on disk, for a relatively coarse source space (even after reducing the sensor level data via PCA). One option would be to now turn towards high performance computing and parallelise this computation, but we instead opt to reduce the number of covariance matrices to a more manageable number, P, on the order of a few hundred.

We do this by running K-means clustering (Hartigan and Wong, 1979) on the dynamic coefficients, α_t. The centroids of the learnt clusters of the alphas are used to construct the P covariance matrices:

$$\mathbf{C}_Y^t \approx \mathbf{C}_Y^p = \sum_i^K \alpha_i^p \beta_i. \tag{4.6}$$

Thus, at timepoint t, we use a look-up table to find which centroid of the k-means-clustered alpha terms we should use to approximate the instantaneous dynamic covariance. Figure 4.2 depicts this concept for a fictional set of α_t values.

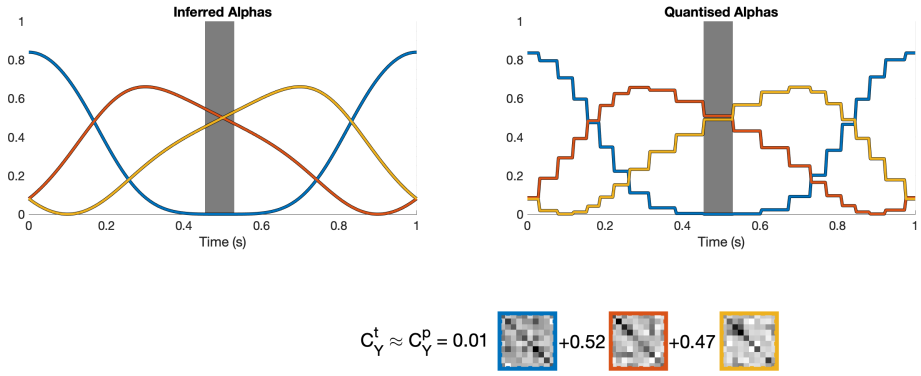


Figure 4.2: Reducing the dimensionality of our parameter space via clustering. The (possibly infinite) number of combinations of the inferred α_t timeseries (left) is quelled into a fixed number of clusters (right) in order to approximate the dynamic coefficients (see shaded area). The learnt centroids are then used to approximate the linear weighted sum by means of a look-up table at timepoint t (bottom).

Simulations

The idea of using time-varying source localisation weights is an attractive one. To carry out our due diligence, we must first ask when we should and when we should not adopt such a method.

Woolrich et al. (2013) used a HMM to estimate different covariance matrices from non-contiguous timepoints. However, in their simulation, they only investigated the effect of being dynamic in a relatively low biological SNR setting, where the signal of interest had a variance 10x smaller than that of the (relatively distant) neighbouring dynamic confound sources. Moreover, the authors did not set out to exhaustively quantify the full effects of being dynamic within the context of a beamformer, but instead opted to present a handful of correlation scores (between the simulated and reconstructed timeseries) for the static and dynamic approaches. The authors did however indicate that the benefits of using multiple covariance matrices increased as the measurement noise increased. At the other end of the data spectrum, Barratt et al. (2018) demonstrated that the advantages of dynamic source localisation start to dissipate as the measurement noise becomes very high.

In these simulations, we investigated the advantages of using numerous reconstruction weights across a range of biological and measurement SNRs. We achieved this by permuting the location of the vertex of interest around the cortex, thus achieving different realisations where the biological SNR ranged from being inherently low (when the source is deep, at the bottom of a sulcus or is radial) to inherently high. We repeated these simulations across different brain regions, to see what effect (if any) the source positions had on results.

Note that the purpose of these simulations is solely to demonstrate where a more sophisticated approach to beamformer source localisation may be appropriate, in the idealised situation where we know what the empirical time-varying covariance is.

Setup

The head position from a previous MEG experiment was used to position a canonical mesh inside of a CTF MEG dewar. We used the Nolte single shell forward model to calculate the lead field.

A source “seed" time course was placed at a vertex on the cortical surface. We put the seed in 4 different brain areas: the left posterior occipital lobe, left primary motor cortex, right pre-frontal cortex and the left anterior temporal lobe. Within each brain area, the exact location of the source of interest was varied upon each instance of the simulation to control for differing lead field magnitudes. We simulated 250 datasets per brain area at each SNR.

Nine dynamic confound sources were placed at random vertices around the brain at each simulation instance, ensuring that the whole of cortex was approximately covered. A tenth source was placed such that it was within 20mm of the vertex of interest. Each confound source comprised of 5 vertices, corresponding to a patch of activity of about 5mm. We ensured that the timeseries of the vertices within each patch had no correlation. Each of these sources turned on (exclusively) for 300ms bursts at random points in time. Both seed and nuisance vertices comprised of white noise band pass filtered to between 1 and 20Hz, with standard deviation equal to 20nAm.

100 additional vertices of small (std = 0.1nAm) amplitude were added at random locations to increase the rank of the source space data. In all our simulations, we synthesised 10 minutes of data using a sampling frequency of 400Hz unless otherwise stated.

The signal to noise ratio at the sensor level was controlled by adding white noise to the clean biological signal after it was projected through the lead field. We used SNRs of -30 to 0 dB in steps of 5dB. To ascertain the performance of TASER in the presence of realistic extracranial biological noise, we added the artefactual components from a previous real ICA result to the (now contaminated) sensor level data. Here we used time courses associated with heartbeat and ocular movements. Finally, we band-pass filtered the resulting dataset between 1 and 48Hz, like we would do with a standard MEG dataset.

We reconstructed the data using a) a static LCMV beamformer and b) with a dynamic beamformer, where we used the known ground truth time-series of the dynamic sources to construct 10 different covariance matrices.

Figure 4.3 shows one instance of our simulation setup and summarises the philosophy of a dynamic localisation regime.

We ran 3 different versions of the simulations:

1. In the first simulation, the location of the nuisance vertices was controlled such that at least one source was within 20mm of the vertex of interest.
2. We repeated simulation 1 in the motor cortex but this time added data from a real empty room recording. We chose 10 continuous minutes of data at random from one 30 minute empty room recording. We applied the same pre-processing steps to the ERN data (downsampling to 400Hz and band-pass filtering, but with an additional stop-band filter between 48 and 52Hz).
3. In the third simulation, we varied the duration of the data set length. We synthesised data at 2 minutes, 4 minutes and 6 minutes.

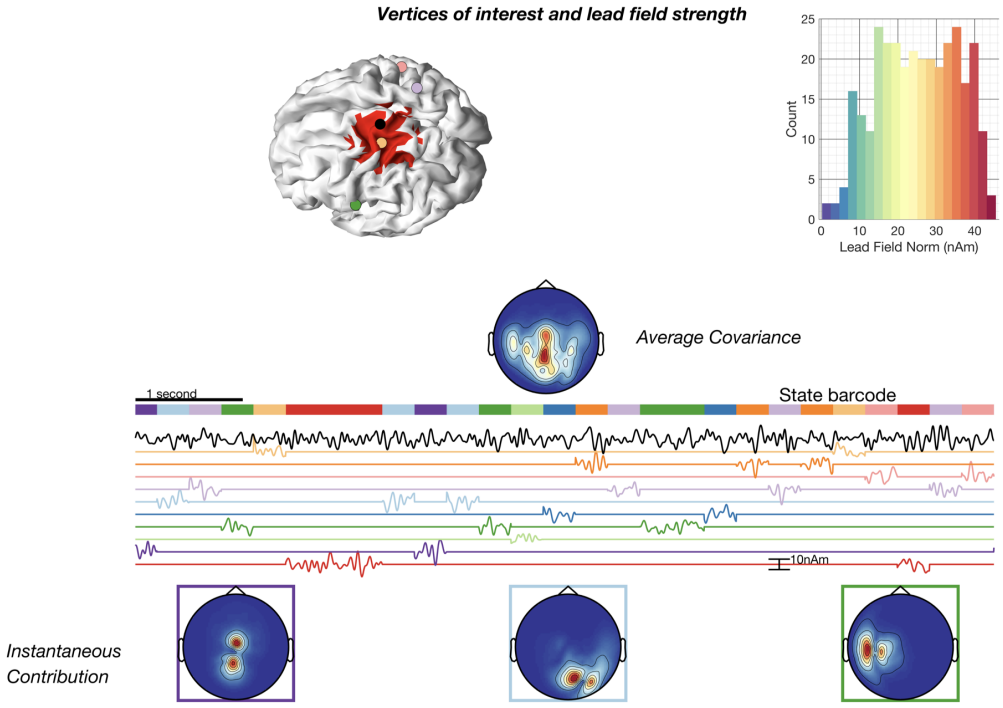


Figure 4.3: The signal of interest (black sphere) is placed at a vertex within a particular region upon each instance of the algorithm. In this example the seed is placed at one of the red vertices in left motor cortex. The spectral plot shows the inherent range of sensitivities to sources within the ROI, due to both source depth and orientation. 10 confound sources are placed at random locations around cortex (coloured spheres) which burst on and off for 300ms; an example source space timeseries for these dynamic vertices and the vertex of interest are shown. Importantly, if one collapses the covariance estimation over all time, the instantaneous covariance estimation is degraded and deviates from the ground truth. The bottom panel shows the covariance contribution for 3 of the individual sources. This has implications for source space results.

Results

Simulation 1

Figure 4.4 shows how the correlation between the ground truth and reconstructed timeseries varies as a function of both brain region and sensor level SNR for both the static (orange line) and dynamic beamformer (blue line). Results have been ordered in terms of reconstruction accuracy to aid visualisation. In all cases, when the noise is very high, the beamformer reconstructs the timeseries poorly. As the quality of data increases, we begin to see differences between the two approaches. In instances that the static beamformer fails to faithfully reconstruct the timeseries, we see an improvement in correlation score when switching to the dynamic regime. This is most noticeable when correlations are <0.7; such improvements begin to dissipate as the static beamformer's performance increases.

Interestingly, in the brain areas where there is good SNR (left occipital and motor cortices here), when the data become very clean, the static approach outperforms any attempt to account for covariance dynamics. However, it should be stressed that we only see this switch at very high correlations ($r < 0.95$). In other words, the loss in reconstruction accuracy is very small and is only likely to impact the most superficial vertices (see Figure 4.5), and only at very high SNR levels. In contrast, for brain areas with lower spatial coverage (e.g. parietal and prefrontal

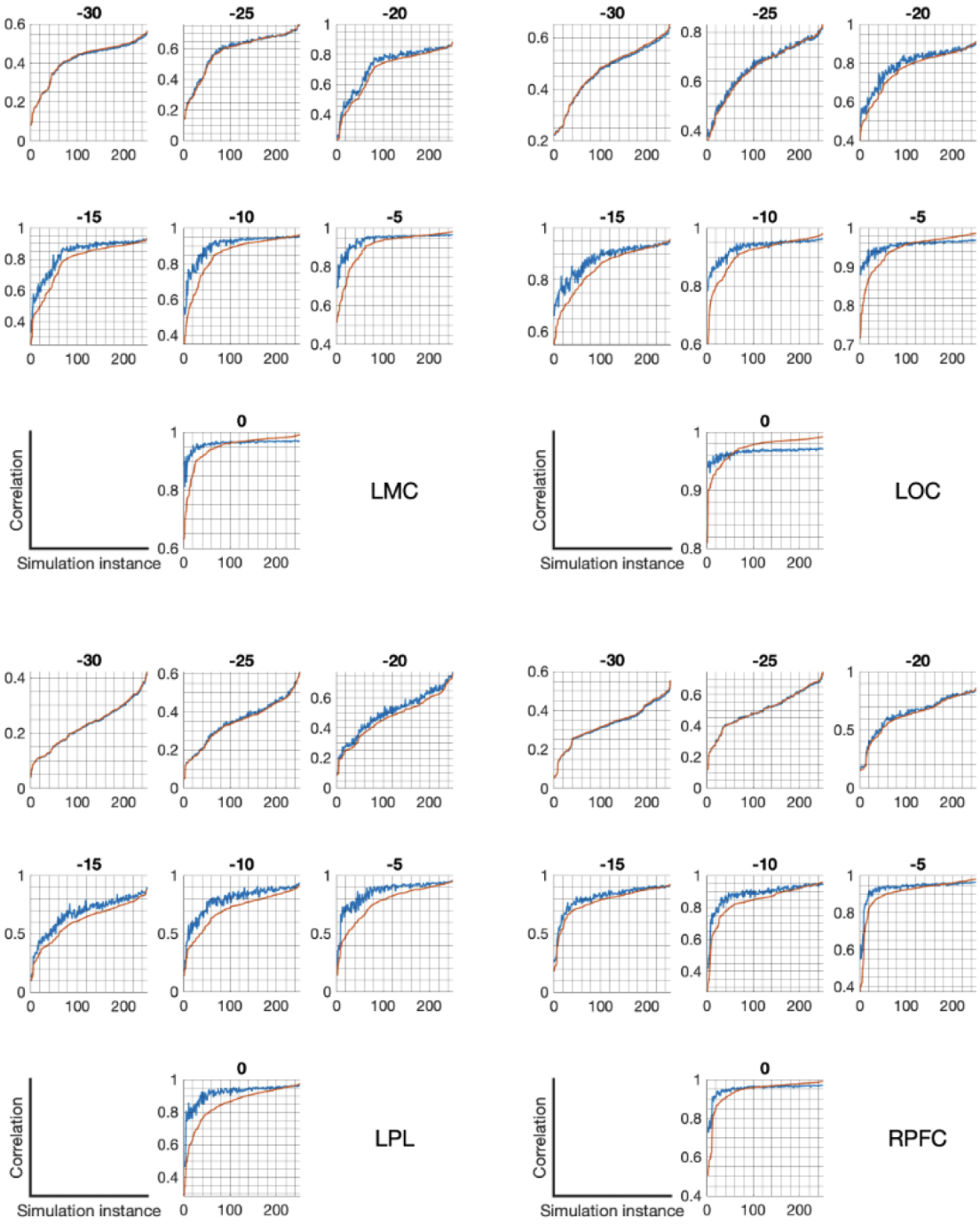


Figure 4.4: Correlation scores across the brain at different SNRs. The static approach is shown in orange, the dynamic in blue. As signal quality improves, we begin to see the benefits of using dynamic source reconstruction, with regions of the brain with lower spatial coverage standing to benefit the most. LMC = left motor cortex, LOC = left occipital cortex, LPL = left parietal lobe, RPFC = right prefrontal cortex.

cortices), we see a continual improvement in reconstruction accuracy, with the largest improvements again seen at the highest sensor level SNRs.

Figure 4.5 shows the impact of lead field strength on correlation change (dynamic minus static) across SNRs, with results pooled across the 4 brain areas. At the lowest SNRs, there is very little difference between the two approaches. However,

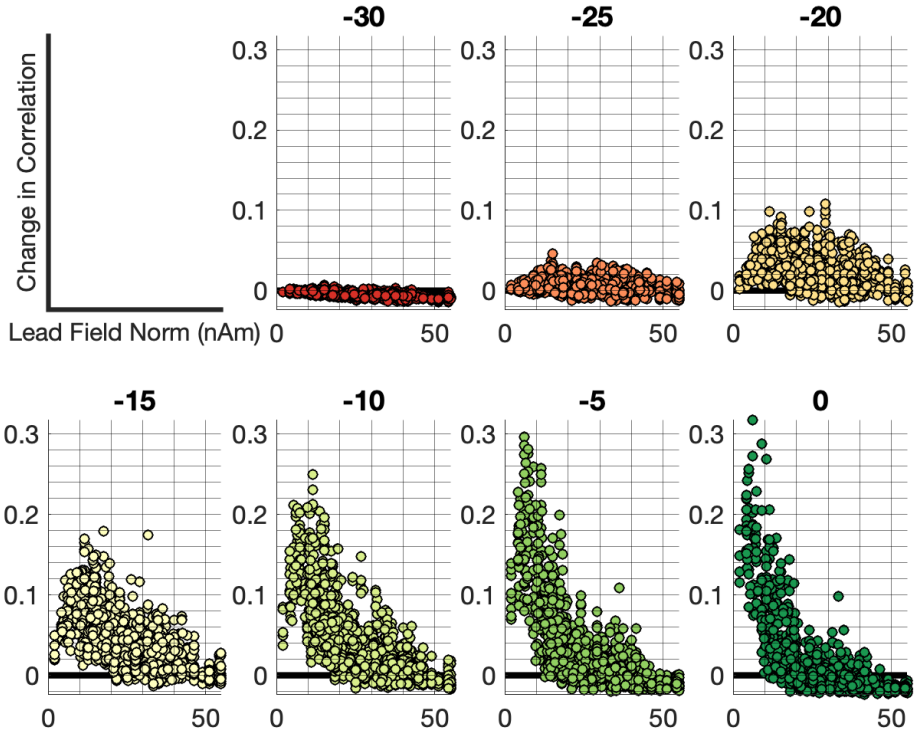


Figure 4.5: Lead field strength appears to determine dynamic source localisation performance. When data are of sufficient quality at the sensor level, there is a marked increase in reconstruction accuracy for lower SNR vertices within the brain. Importantly, any decreases in performance are very small (c. 0.02) and only happen when the data is extremely clean.

as the data become cleaner, we begin to see a clear relationship between the lead field norm and the change in correlation; vertices with inherently lower sensor level representation stand to gain more from a dynamic source localisation regime, with improvements over 0.3 in the most extreme cases. At these extremities of data quality, we see a small (0.02) decrease in correlation for the strongest vertices.

Our results in Figure 4.6 show that there doesn't appear to be any statistical relationship (Pearson correlation of 0.03 (p=0.2401) at SNR = 0dB) between the correlation change and smallest distance between source and nuisance vertex.

Simulation 2

Figure 4.7 shows the same relationships as in Simulation 1, but this time for when real empty room recordings were used with a seed in left motor cortex. We see the same patterns emerging as before. Figure 4.8 shows an example instance of the ground truth and reconstructed timeseries.

Simulation 3

Figure 4.9 shows the impact of decreasing the number of data points available for covariance estimation. These findings are in accordance with previously published literature (Brookes et al., 2008), with reconstruction accuracies dropping as fewer

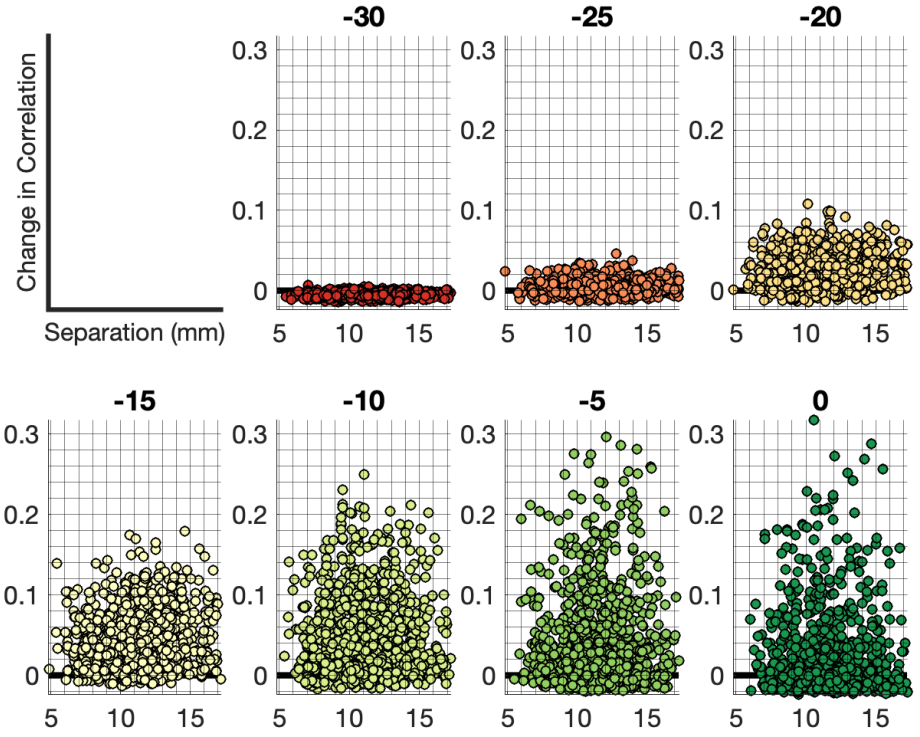


Figure 4.6: Separation between the vertex of interest and nearest dynamic confound source does not appear to impact correlation scores of source localised activity.

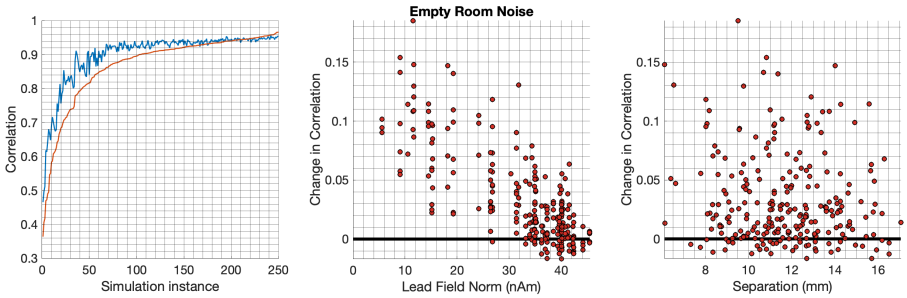


Figure 4.7: Correlation (change) as a function of lead field strength and vertex separation when real empty room recordings were used instead of white noise in the motor cortex case. We see similar patterns as shown in previous Figures.

timepoints are used in the covariance calculation. The benefits of using a dynamic beamformer return as the length of the dataset increases above 4 minutes.

Discussion

Our simulations show that, if one can confidently estimate different covariance matrices (and thus reconstruction weights) for different periods of time, then there is little (to no) risk in adopting a more ambitious approach to beamformer source localisation. Indeed, we only saw a drop in Pearson correlation when data became very clean and the static beamformer achieved scores of > 0.95. In such cases, the dynamic beamformer's drop in accuracy was marginal (0.02).

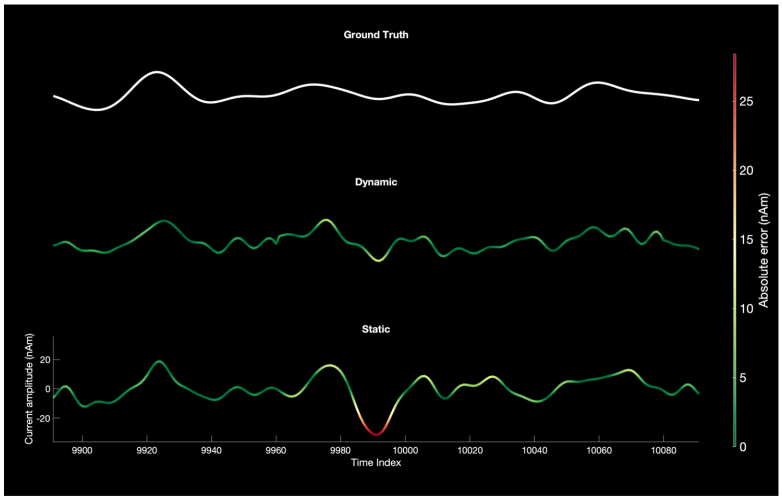


Figure 4.8: Example result for the motor cortex empty room recording (Simulation 2). Top: ground truth timeseries; Middle: dynamic result, where the underlying time-varying covariance matrix was used; Bottom: reconstructed time-series when the covariance matrix was collapsed over the entire experimental window. All data are plotted on the same scale, with absolute error between the reconstructed and ground truth timeseries encoded with colour to aid visualisation. Note the large dip in accuracy in the static approach around the middle of the window, and other notable deviations elsewhere (yellow colours).

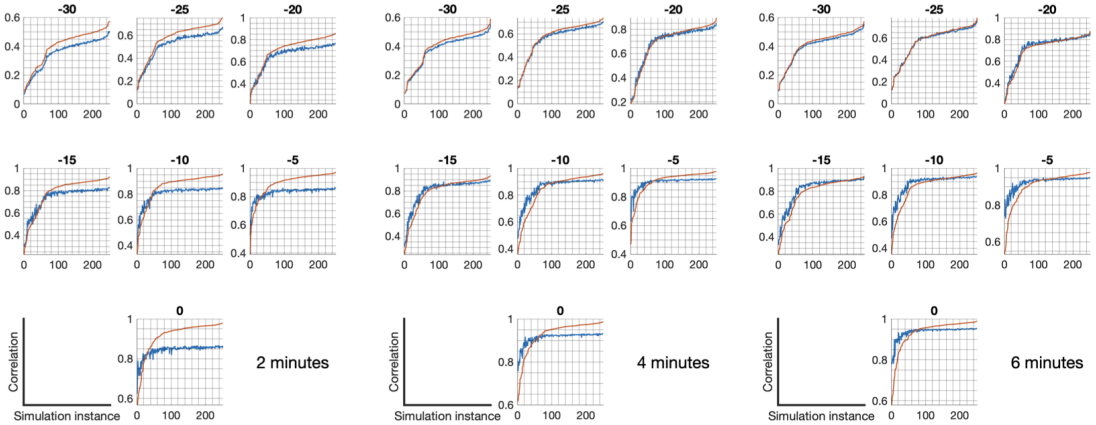


Figure 4.9: The impact of dataset length on the correlation accuracy in the motor cortex case.

One of the advantages of the beamformer Equation is that every vertex in the brain is considered independent of one another. One could therefore envisage a hybrid approach between using a static covariance matrix for some vertices, and a dynamic set of weights in areas of small lead field norms.

The results presented here indicate that dynamic beamformers may help researchers to investigate brain regions which are classically considered “hard" to study with MEG. This includes deeper areas of cortex, or even subcortical regions. The reader should also be reminded that the lead field is a function of source orientation, with more radial brain areas being less sensitive to MEG sensors. We therefore hypothesise that a dynamic beamformer may do a better job at reconstructing more spherical deep brain structures such as the thalamus or amygdala.

Our findings in this Chapter start to become intuitive if one is to stare at Equation

(2.9) for long enough. When the norm of the lead field is large with respect to the elements of the covariance matrix, then it is likely that the forward model will swamp out any subtle changes in the data covariance. Conversely, when the lead field vector is relatively weak, then there is scope for any changes in covariance to influence the calculation of the reconstruction weights.

Source space functional connectivity (FC) studies are directly affected by source reconstruction (in)accuracies, where faithfully reconstructing the vertex timeseries is crucial. Put another way, a FC expert is likely to care more about Pearson correlations than other source localisation accuracy metric. Having focussed on correlation here, our results suggest that a dynamic beamformer would do a better job at assessing the FC between one low SNR region (e.g. auditory cortex) and any other.

We must stress that the usual rules of data covariance estimation apply (Brookes 2008); simulation 3 highlights the drop in reconstruction accuracy when the dataset becomes too short. However, we do start to see improvements from datasets as short as 4 minutes, albeit here such increases in accuracy are smaller. This is good news for the MEG community, where a typical resting state paradigm will last only 5 minutes. A possibly exciting future avenue of research would be to repeat the simulations presented here for a smaller number of covariance matrices, where this degradation to due data set length should start to reduce.

One potential criticism of our simulation methods could be the use of just one empty room noise recording, and the fact that we took random (continuous) segments from this. An alternative approach would have been to fit an autoregressive (AR) model to the empty room recordings and sampled synthetic datasets *ad infinitum* from the learnt AR coefficients. However, this would have introduced uncertainties about whether or not the synthetic empty room noise had realistic spectral and spatial properties.

Being simulations, the forward model was “perfect" in this work. One important future piece of work should investigate how quickly these theoretical gains disappear with realistic coregistration error, a point we shall return to in later chapters.

We placed the seed in motor cortex for the results obtained in Figure 4.7. Stereotypically this has very good coverage in cryogenic MEG experiments, and so by extrapolating our results from Figures 4.4 and 4.5, we would expect to see a more marked improvement for weaker brain areas.

Above all, we must reiterate the fact that this investigation only looked at what improvements may be gained when the dynamic empirical covariance is known without error. The real challenge now lies at the feet of the data analyst who must make dynamic estimations in the presence of a real experiment with an unknown ground truth. The simulated timeseries of the confound sources here were unrealistically simple (i.e. categorical). Future inference methods should be able to take into account non-categorical covariance dynamics if these plausible improvements in LCMV beamformers are to be realised. One such method was introduced in Chapter 3.

To conclude, in this section, we have shown that:

- Dynamic beamformers can improve source localisation correlations by more than 30%
- Brain regions which are less well covered are more likely to benefit from temporally adaptive methods
- Nearest neighbour distance does not impact changes in source localisation correlation accuracy for dynamic versus static methods

- There is little risk to introducing dynamic reconstruction weights, provided that covariance matrices are accurately estimated.

Conclusion

In this chapter we have introduced the TASER philosophy to source reconstruction and shown how it can be understood as a temporal extension to existing inverse solutions. We have also identified the regimes under which a more sophisticated dynamic approach to source localisation may prove beneficial within the context of a beamformer. We have presented a non-trivial relationship between the lead field magnitude, the sensor level SNR and the reconstruction accuracy. Importantly, such parameters are either deterministic or easy to estimate, respectively, ahead of carrying out source localisation. We therefore believe that these results will aid future researchers in deciding whether or not to use non-stationary representations of the data covariance in their source localisation pipeline.

Our simulations show that temporal reconstruction accuracy increases when dynamics are incorporated for weakly represented brain regions. Previous work has shown that dynamic localisation is also necessary when trying to carry out MEG studies with high spatial precision.



5 The Pinky and The Brain

Introduction

A practical problem faced by any novel source analysis tool is the absence of a ground truth solution when tackling real data for the first time. Commonly, researchers will first turn to a well classified and understood electrophysiological response across space and time and both qualitatively and quantitatively assess the performance of their new pipeline, making comparisons with established methods such as minimum norm and the beamformer.

We adopt such an approach here, applying the beamformer flavour of TASER to a previously published somato/motortopic study, using the known organisation of the human homunculus as a proxy for a ground truth. We also include exemplar implementations of TAMSP to demonstrate the technique for a single session.

The structure of this chapter is now outlined. We begin by outlining the motivations and key findings of the original work. The acquisition, experimental paradigm and the analysis pipeline is then detailed. Sensor and source level results are discussed in detail for one subject across different dynamic techniques before a quantitative analysis over all subjects and sessions is presented. We end the chapter with a discussion of the results and methods introduced here.

Background

For decades, MEG has been thought of as a technique with inherently poor spatial resolution and is often discounted as a functional neuroimaging method when the goal of a study is to localise a source of interest with high spatial resolution. This is in spite the fact that numerous studies, both numerical and experimental in nature, have shown this to not be the case (see, for example, Hillebrand and Barnes (2005)).

Much of this previous work includes somatotopic mapping. In an early *tour de force*, Yang et al. (1993) used a 37 channel MEG system to map the evoked response to the stimulation of 66 different body sites. Nakamura et al. (1998) repeated this feat using an air puff stimulator to map evoked responses to stimulation of various different body parts on the right-hand side of participant, whilst Fisher et al. (2004) used a whole-head modern MEG scanner to spatially disambiguate the clunis from the cubitus. Momentarily turning our attention away from the M/EEG literature, recently Roux et al. (2018) reported the findings of electrostimulation to

BA1 in awake and intact patients undergoing lesion removal. Patients' cortices were stimulated with small currents ($\sim 1.5mA$) that were increased in magnitude until a physical sensation ("tingling") was reported. Researchers found that the sensation in individual fingers was tightly localised to areas on the order of $3mm^2$ on the surface of cortex, a finding whose relevance to this chapter shall soon become clear. Finally, to highlight just one study, O'Neill et al. (2020) used 7T fMRI to create probabilistic maps of individual digits arising from stimulation of both hands. Interestingly, they showed large heterogeneity in the localisation of digit 5 for N=22 right-handed participants.

This body of work demonstrates the utility and popularity of adopting a paradigm as simple as asking subjects to tap their fingers. Historically however, MEG and EEG researchers have tended to map the averaged (time-locked) response to such digit activation and have shied away from studying spontaneous or non-phase locked activity around movement onset and offset. This is for two primary reasons. First, by collating hundreds or thousands of trials in a very short period of time[1], experimenters are able to dramatically increase the SNR of averaged responses and are even able to image radial and deep patches of cortex where spatial coverage is low (e.g. Nakamura et al. (1998)). Second, the source imaging of spontaneous rhythms is a much newer endeavour in comparison to equivalent dipole fitting.

Dataset Synopsis

Barratt et al. (2018) changed this trend recently when they studied the spatial topology of non-time-locked oscillations in the beta band in human motor cortex. Using subject-specific foam headcasts, they mapped the brain's response to movement cessation, imaging the post-movement beta rebound (PMBR) resulting from index and pinky finger activation. Their results claim to spatially disambiguate the brain's response to these different finger movements. These findings were further verified by a supplementary fMRI study using an analogous paradigm with the same subjects. The authors argued extensively that the MEG result was only possible due to the accurate forward modelling afforded by the foam headcasts and, relevant to this work, that the accuracy required to delineate the digits (in space) was achieved by manually constructing different beamformer weights for each finger.

We attempt to reproduce these findings using static and TASER methods in this chapter. In the temporally adaptive approaches, we remove the need to "hand-draw" covariance windows for different experimental conditions. Instead, we let the data drive us towards the optimal combination of elements which sum to give a dynamic picture of the covariance in a completely unsupervised manner, a task which would be almost impossible to achieve at the behest of the experimenter.

Methods

The raw MEG data from the original publication were kindly shared with us. We attempted to keep all preprocessing settings the same throughout the work to facilitate a direct comparison between dynamic reconstruction and the static methods presented in the original paper. Any differences in the analysis procedure are highlighted and justified below.

[1] Such short inter-trial periods are not feasible when studying induced oscillations, as post-event responses decay on slow (second) timescales - see later in this chapter.

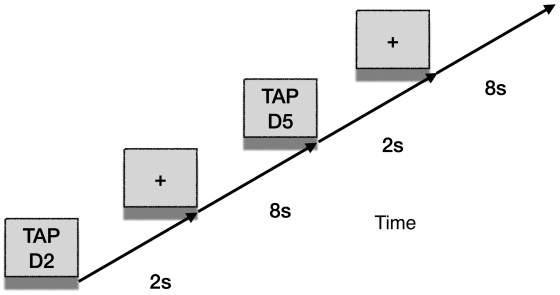


Figure 5.1: Paradigm for the MEG component of the experiment. Adapted from Barratt et al. (2018).

Ethical Consent

Three healthy adults (1 female - age 35 years; 2 male - aged 26 and 24 years) took part in the study, having given full written consent. The study was approved by the University of Nottingham Medical School Research Ethics Committee. Each subject was scanned 8 times in the MEG system across 2-4 days. Functional MRI scans were also acquired, but we did not analyse these data in this follow up study. In all that follows, we will only detail the analysis of the MEG data.

Experimental Paradigm

The experimental paradigm was designed to optimally spatially distinguish PMBR on digits 1 and 5. Essentially, this boils down to including a sufficiently long baseline between trials such that any PMBR decays and returns to baseline (Fry et al., 2016). A visual cue instructed subjects to “tap" (execute flexion and extension) either D2 or D5 against a surface at a frequency of approximately $2-4Hz$; no EMG data were acquired so this was not monitored explicitly. Each trial lasted 20s and comprised 2s of tapping D2, followed by 8s rest, and then 2s of tapping D5, followed by a further 8s rest with 44 trials collected. The entire MEG experiment was then repeated 8 times in each subject. See Figure 5.1.

Foam Headcast Construction

To generate the headcast, an anatomical MRI of the subject’s head was acquired using a 3T Philips Ingenia scanner running a 3DTFE sequence with resolution of $1mm^3$. Sequence parameters ($TR/TE = 4.5ms/1.97ms$; $FOV(AP, FH, RL) = 256x256x183$; $BW = 775Hz$) were optimised to provide a high bandwidth/voxel and hence minimise spatial distortion due to susceptibility effects around the face and scalp. A 3D mesh representing the outer surface of the head and face was extracted from the anatomical MRI. This 3D model of the subject’s head was placed inside a virtual realisation of the MEG helmet. The head-model was augmented with spacing elements, to optimise the distance between the scalp and sensors. In addition, a 3D representation of three head localisation coils, used for localising head position in the MEG scanner was added, with coils placed at the nasion and left and right preauricular points. This 3D digital representation of the head surface, spacers, and localisation coils was 3D printed and placed inside a manufacturer provided replica of the MEG helmet. Finally, liquid resin was poured into the negative space between the surfaces. This expands and sets, resulting in flexible foam headcasts which can also house the localising coils.

MEG Acquisition

All MEG data were acquired using a 275-channel whole head CTF system (MISL, Coquitlam, Canada) operating in third order gradiometer configuration at a sampling

frequency of 600Hz. Three localisation coils were placed in the headcast and were energised continuously during the scan, facilitating coregistration. Visual cues for the MEG paradigm were presented via projection through a waveguide onto a back projection screen placed $\sim 40cm$ in front of the subject.

Preprocessing

All MEG data were inspected visually by the original experimenters. Any trials containing excessive interference (e.g. due to the magnetomyogram) were removed - this resulted in the reduction of the trial count to 42 ± 2 in Subject 1; 41 ± 3 in Subject 2 and 42 ± 2 in Subject 3 (mean and standard deviation across 8 runs in each subject). The mean (across 8 runs) maximum movement for subjects 1, 2 and 3 was $0.3mm$, $1.2mm$ and $0.3mm$ respectively. We did not carry out any further data cleaning (e.g. ICA) in the follow up study, and kept all trials which were provided to us in the original data files.

All preprocessing was carried out using FieldTrip (http://fieldtriptoolbox.org, (Oostenveld et al., 2011)). We bandpass filtered the data between 13 and 30Hz and used a discrete Fourier Transform at 50, 100 and 150Hz to remove line noise and the harmonics. To speed up inference, we downsampled our data to 200Hz (unlike the original work where data were kept at 600Hz).

To facilitate the construction of an accurate headmodel, the anatomical MRI for each subject was segmented into grey matter, white matter and CSF using SPM8. Anatomical data were coregistered with the MEG sensor array at the time of the experiment, removing the need for any additional coregistration steps. A Nolte Single Shell (Nolte, 2003) volume conduction headmodel was used for each subject and a 6mm isotropic grid (bounded by the inner skull) defined the discretised source space. The lead field for three elemental currents was calculated at each of these locations; the rank of the resulting vector at each voxel was reduced to two to eliminate the weakly contributing radial component (Ahlfors et al., 2010; Hämäläinen et al., 1993). Lead fields were not normalised.

Analysis

Statistical Procedure

We reproduced the results published by Barratt et al. (2018) by constructing images showing power changes in the “active" and “baseline" windows. We used the same timepoints (with respect to the start of each epoch) for these windows as in the original work: 2.5 to 4s for digit 1 with a baseline of 7.5 to 9s, and 12.5 to 14s and a baseline of 17.5 to 19s for digit 5. In the original work a pseudo T-statistic was reported. Generally, this takes the form of

$$T_{pseudo} = \frac{\mathbf{w}^T\mathbf{C}_A\mathbf{w} - \mathbf{w}^T\mathbf{C}_C\mathbf{w}}{2\mathbf{w}^T\mathbf{w}}, \tag{5.1}$$

where $\mathbf{w}$ is the weight vector for the i^{th} voxel/vertex and $\mathbf{C}_A$ and $\mathbf{C}_C$ are the covariance matrices calculated in the “active" and “control" windows, respectively.

Despite its simplicity (or perhaps, because of it), this Equation is very powerful and popular within the beamforming community (Vrba and Robinson, 2002; Brookes et al., 2010). However, it is immediately apparent that it is unsuitable when either the weights vector, and/or the covariance matrices change over time. For this reason, we first source reconstructed the data across the whole brain, before calculating the power in the active and control windows (sum of squares) for each voxel. This accounted for the time-varying nature of the covariance in the active and control

windows. We then performed a non-parametric independent samples T-statistic on these vectors of power values, $\rho \in \mathcal{R}^{N_{trials} \times 1}$. Statistics were corrected for multiple comparisons using cluster-based randomisation ($N_{permutations} = 1000$) as implemented in FieldTrip (Maris and Oostenveld, 2007).

Beamforming

We used various different stationary and nonstationary forms of the LCMV beamformer. We detail these below. All analysis was done on a session-by-session basis. Note that we kept the regularisation parameter constant for each of the different covariance matrices (1% of the mean eigenvalue of the covariance matrix). We kept the first 50 principal components of the pre-processed sensor level data when running the HMM and VRAD inference. We therefore used the first 48 singular values of the covariance matrix when inverting it.

1. **“Static" Beamformer**
   All data across the whole experimental window (20s x 44 trials) were used to calculate the covariance matrix then used in the construction of the beamformer weights, giving $N_w = 1$ weights matrix.

2. **“Manually Refined" Beamformer**
   Covariance matrices were constructed from data segregated according to their experimental condition. Specifically, one covariance matrix was calculated from the first 10 seconds of the epoch (pertaining to index finger trials), and a second matrix was defined from the data in the second half of the epoch. Data were source localised separately using two different sets of weights for each of these conditions, $N_w = 2$.

3. **HMM Beamformer**
   We ran a HMM with a (zero-meaned) Gaussian observation model separately for each of the 8 sessions and 3 subjects on the pre-processed sensor level data. Prior to inference, we reduced the dimensionality of the timeseries by keeping the top 50 principal components. This typically resulted in 97% of the variance being explained. Importantly, we also normalised the data such that they had zero mean and unit variance on a channel-wise basis. Previous work has showed that typically a large number of HMM states were required to see sufficient differences between a static and dynamic approach to source localisation (Woolrich et al., 2013). We therefore separately inferred models with $K = 11$ and $K = 20$ states. Upon convergence, the learnt covariances were then projected back out to the native $N_{channels} \times N_{channels}$ space and rescaled to their original variance using the Hadamard product. These covariance matrices defined $N_w = 11$ and $N_w = 20$ weights matrices. We then used the Viterbi path from the learnt HMM state time course to determine which set of reconstruction weights to use at what points in time.

4. **VRAD Beamformer**
   Finally, we used the novel inference framework to estimate the instantaneous covariance across the whole experiment (using the same data dimensionality reduction steps detailed above for the HMM-powered dynamic inference), in theory giving $N_w = T$ reconstruction weights matrices. To reduce the computational burden and redundancy in the combination of different basis functions, we used K-means clustering to learn 400 centroids of the alpha coefficients, thus giving $N_w = 400$ source reconstruction weights (see Figure 4.2).

   We inferred 20 covariance basis functions. These were initialised with “noisy" estimates by taking 20 random 200ms segments of the continuous dataset and calculating the covariance from these windows (c.f. Figure 3.5). Inference was run for 500 epochs (sufficiently long for convergence), with a KL-annealing period for the first 300 epochs. We adopted a hyperbolic tangent annealing

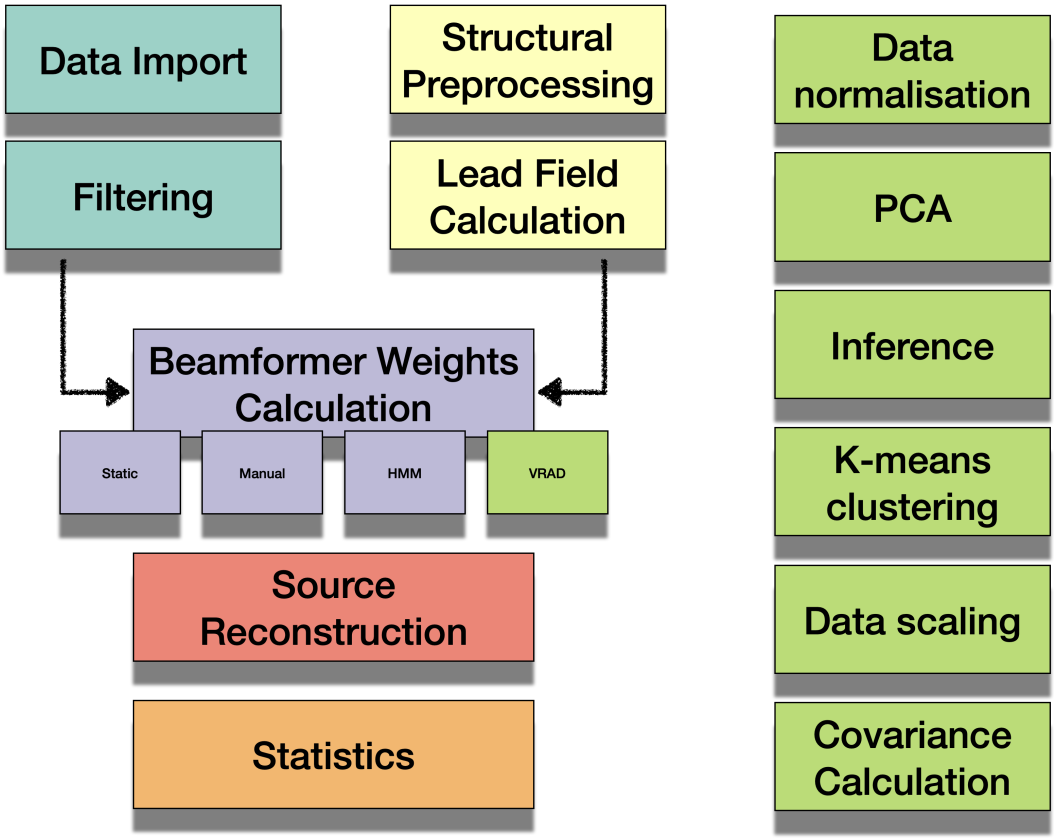


Figure 5.2: Analysis pipeline employed in this chapter. Beamformer weights were calculated (purple) after MEG (teal) and structural (yellow) data were pre-processed. The sub steps of the VRAD inference are shown on the right (green). These steps were carried out for the HMM, minus the K-means clustering.

profile with a sharpness of 10 (see Equation (3.24)). In this work we used single layer LSTM networks for both the inference and model, with 64 and 96 units, respectively. The inference network used bidirectional RNNs. The sequence length was set to 200 time points (1 second's worth of data in this study) and the batch size equal to 64 (giving approximately 14 gradient updates per training epoch). We did not use dropout between the layers. The outputs of the RNN and theta activation layer were normalised (independently, across each example i.e. across features and not batches) to improve numerical stability (Ba et al., 2016). The learning rate was set to $1e{-}3$ and the alpha temperature was set to 2.

Statistical maps were non-linearly transformed to MNI space with SPM8 after being constructed in MEG head coordinates (Ashburner and Friston, 2005). The entire methods pipeline for the beamformer implementation is summarised in Figure 5.2. Note that the example TAMSP results presented here were subject to a slightly different pipeline. These are provided as an appendix to this Chapter.

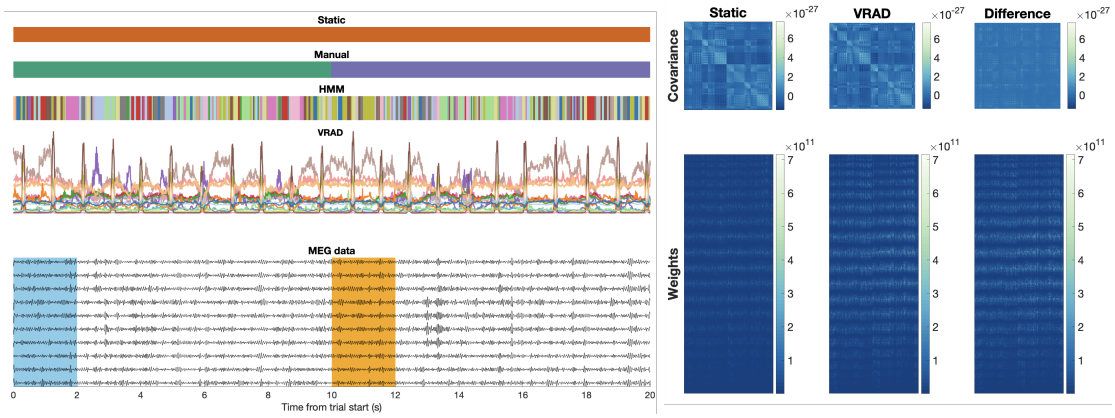


Figure 5.3: Different approaches to temporal modelling give different representations of the data. Left: schematics and inferred representations of the covariance dynamics. The four different modelling options are shown; one covariance matrix for the entirety of the experiment or trial (orange); covariance matrices manually constructed according to the task structure (green/purple); a 20-state HMM (state barcode is shown with different colours) and the VRAD approach where mixing coefficients (α_t, shown as traces) give an estimation to the instantaneous covariance. The corresponding MEG data is shown below for reference; the blue rectangle highlights the period where the index finger was tapped, whilst orange pertains to pinky finger motion. MEG data from the first trial are shown. Manually segregating the data into different experimental conditions oversimplifies the reality of (co)variance non-stationarity, whilst the HMM and VRAD give fluid representations of the changing properties of data at the millisecond timescale. Right: these different approaches lead to different LCMV beamformer weights being derived. Covariance matrices derived from the static and for one timepoint in the VRAD approaches are shown, alongside their difference. The absolute beamformer weights (nAm/fT) and their difference is also shown.

Results

Different temporal models give different perspectives on the same data

Figure 5.3 (left) shows how different modelling assumptions give different perspectives on the latent covariance dynamics. Real inference results are shown for one trial of the dataset alongside 10 channels of sensor level MEG recordings. The top two panels depict the windows used in conventional approaches when beamforming, which are less temporally adaptive than the observed neural data. In contrast, data driven approaches to modelling the covariance give a rapid, transient description of the latent dynamics. As expected, HMMs give a very categorical description of the data (shown here as a state barcode), whilst the α_t coefficients learnt by VRAD show a more temporally continuous picture of the changes in covariance. This has implications for source localisation - the right hand column shows the estimates for the $N_C \times N_C$ sensor level covariance matrix (top) and the corresponding $N_d \times N_C$ beamformer weights (bottom) for the cases of a static and VRAD-powered dynamic beamformer. The difference in the approaches is also shown. Absolute values of the weights are displayed to ease visualisation; the instantaneous VRAD covariance estimation is shown for a randomly selected timepoint.

Incorporating data dynamics leads to more accurate source localisation

A TAMSP Result

Figure 5.4 shows the TAMSP setup and an example result. A comparison is made with SPM's MNE implementation[2]. The T-statistics of the PMBR derived by TASER are much clearer than the classic MNE implementation, with clear disambiguation between the two digits and the peak of the maps in anatomical agreement with the known homunculus. Presented maps are thresholded at the 80th percentile. Panel C shows the patch dynamics for each digit and the (squared) reconstructed signal at the peak vertices. Notice the resemblance to the widely reported motor rebound and the increased power for the respective conditions.

VRAD Beamformer Results

Figure 5.5 shows a source space statistical image of the PMBR for a representative session. The top panel shows the activity reconstructed with a static beamformer. The bottom panel shows the VRAD result. Maps are thresholded at 90% of the maximum T-value. There is substantial spatial overlap of the two maps in the static case. In contrast, the VRAD result not only separates the digits in space, but also recovers a result which is in agreement with the homunculus (see right of Figure 5.6). A map where eigenvalue truncation wasn't used in the matrix inversion is included as an appendix to this chapter for the interested reader.

HMM Beamformer Results

Figure 5.6 (top) shows the conjunction map for the different digits for one exemplar session, overlaid onto the individual's MRI. The teal, orange and magenta colours represent digits 1, 5 and the overlap of their statistics, respectively. Maps were produced by finding the intersection of the underlying PMBR t-statistics thresholded at the 90th percentile.

In all localisation setups, the PMBR is localised to the left postcentral gyrus. When the whole experiment is used to estimate the covariance, there is considerable spatial overlap in the maps for the two digits, making it hard to untangle the true/peak location of the PMBR on the axial and coronal slices. This is demonstrated by the large magenta blob. Interestingly, the sagittal slice appears to indicate that the PMBR associated with index finger activation is more superior to pinky finger motion in contradiction with what is known about the somatosensory homunculus (see cartoon schematic).

A similar result is obtained when covariance windows from the two experimental conditions are constructed, albeit with a slight reduction in the size of the coactive voxels. Taking a dynamic approach to localisation, one begins to see more separation between the two digits when a $K = 11$ state beamformer is used. However, it is not until 20 states are used that we recover the classic homunculus structure, with digit 5 appearing more superior and medial than digit 1. Notice also the much smaller intersection of the two maps for this session.

Assessing Performance

The bottom of Figure 5.6 shows the location of peak rebound statistics in all three MNI directions for the static, $K = 20$ HMM and VRAD-powered beamformers across 8 sessions for the same subject. Small circles indicate each session result, with the mean across sessions indicated by a larger circle. Data are shown on the same

[2]This is an adaptive spatial filter in the sense that the noise and source prior can be scaled by the inference to maximise the VFE, but note that the derived filter remains static in time.

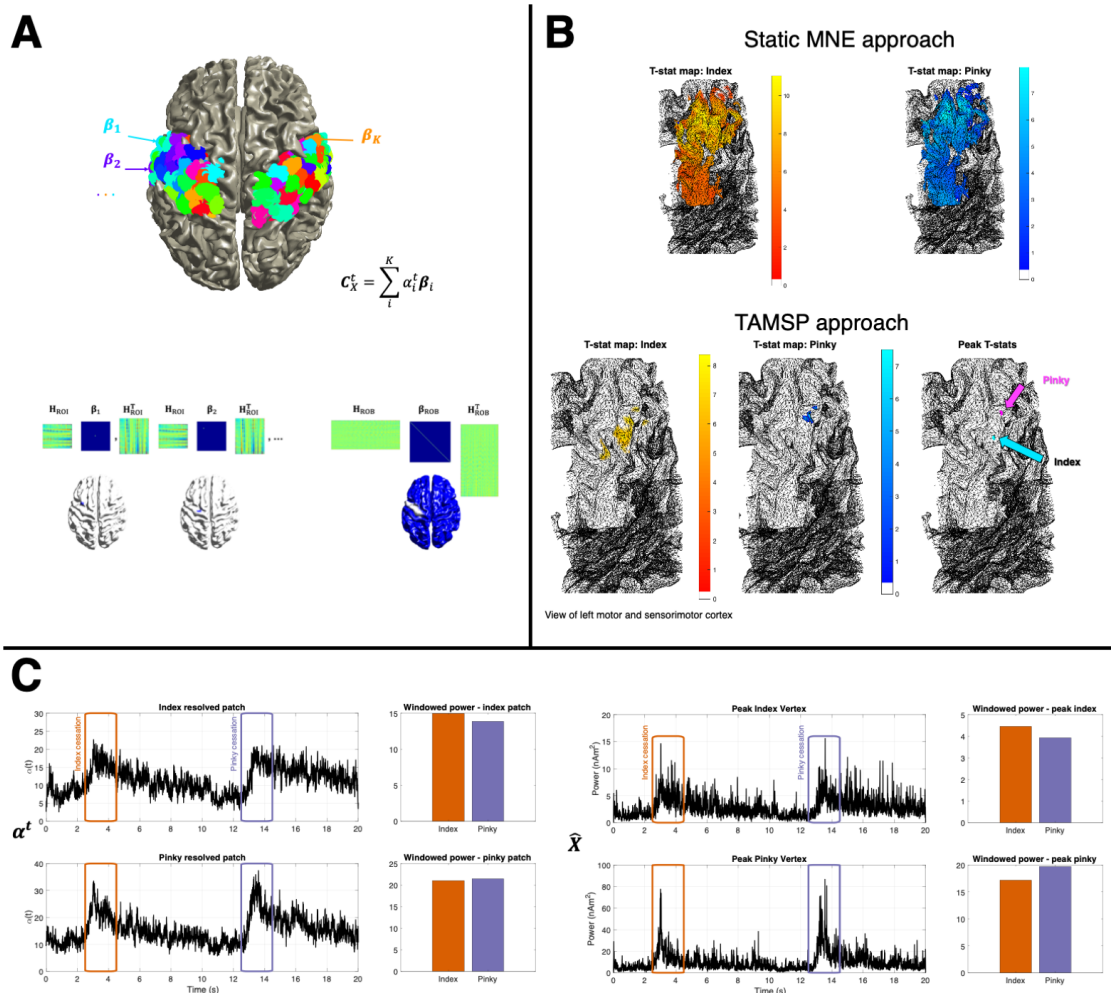


Figure 5.4: An example TAMSP result. Panel A: Priors were packed into motor regions (multicoloured spheres) on a subject's mesh. These patches were projected through the lead field to give sensor priors. The remaining vertices were combined into one MNE-like source prior. Panel B: Comparing the TAMSP result with an adaptive yet static source recon method. Figures show the PMBR T-statistical images. Maps are thresholded at the 80th percentile. SPM's MNE procedure gives no discernible features in the source space. The TASER approach gives a result with two clear peaks which are in agreement with the homunculus. Panel C: inferred mixing coefficients (left) correlate with PMBR dynamics after averaging. Timeseries from the patches containing the peak T-statistic are shown. Bar charts show the increased recruitment of each patch depending on whether a pinky or index flex was executed. Right: Induced source space oscillations at the peak vertices. Again, more power is found in the index patch during an index trial and vice versa. The average of all the reconstructed trials is shown for each vertex.

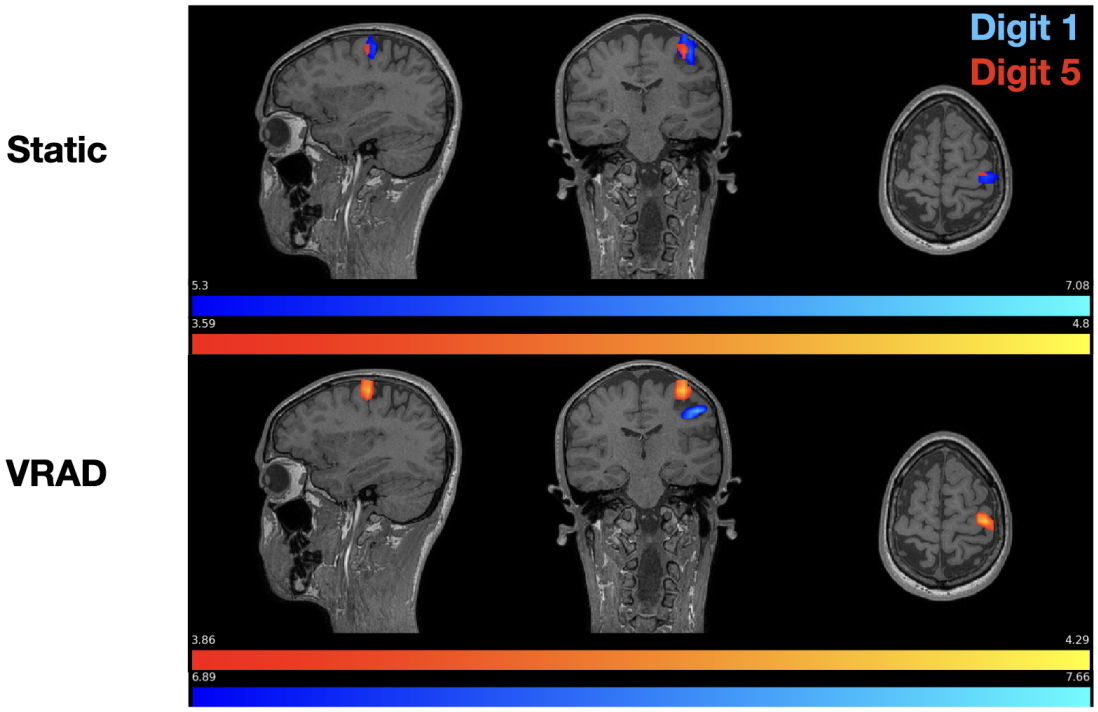


Figure 5.5: TABFER gives more accurate source space images than the classic LCMV method when analysing high-precision MEG data. Maps shown were thresholded at the 90th percentile of the maximum value. The same slices are shown for both approaches.

y scale within each direction. These mean values are tabulated in Table 5.1 for convenience.

We carried out one-sided paired t-tests between the peak t-statistic locations arising from the two digits to test for significant differences in peak maxima. We used the fMRI results to determine which tail of the null distribution to use for each of the orthogonal planes. Asterisks indicate the level of significance: $* = p < 0.05$, $** = p < 0.01$.

As well as using the anatomy to assess our source localisation performance, we also turned to cross-modality functional information to gauge the accuracy of our results. Briefly, we used a probabilistic atlas of individual digits derived from $7T$ fMRI data O'Neill et al. (2020). Atlases were constructed from fMRI data collected in 22 right handed subjects in response to vibrotactile stimulation of the tips of both hands. We took the centre of mass of the probabilistic maps separately for digits 1 and 5 using FSL to derive a point estimate for haemodynamic activation for each digit (Jenkinson et al., 2012). These are shown as horizontal bars on Figure 5.6.

The static beamformer failed to reach statistical significance for all three axes. In the lateral/medial direction the HMM method showed a significant difference ($p = 0.006$) with a mean displacement of $6mm$ whilst the VRAD implementation approached significance ($p = 0.08$). The lack of significance in this plane in the case of the static approach is likely to be down to the session where the statistic peaked in ipsilateral motor cortex, a well known result published elsewhere (Dalal et al., 2008; Huang et al., 2004; Woolrich et al., 2013). However, all methods were in tight concordance with the fMRI results. The quality of the results in this plane is perhaps unsurprising, given the expected clear spatial distinction in the homunculus (see schematic). Significance was not reached in the anterior/posterior plane for any method with the results being at odds with the fMRI statistics. It should be noted

that shifts in this plane correspond to movements within the gyrus, so perhaps it is not too surprising to see discrepancies in this direction. Finally, results were not significant in the MNI Z direction for the static and HMM methods, but were significant for the VRAD implementation, with the digit 5 peak statistic being much closer to the fMRI result. This result, closer to the functionally and anatomically defined ground truths, shows the benefit of accounting for covariance dynamics in a non-categorical manner.

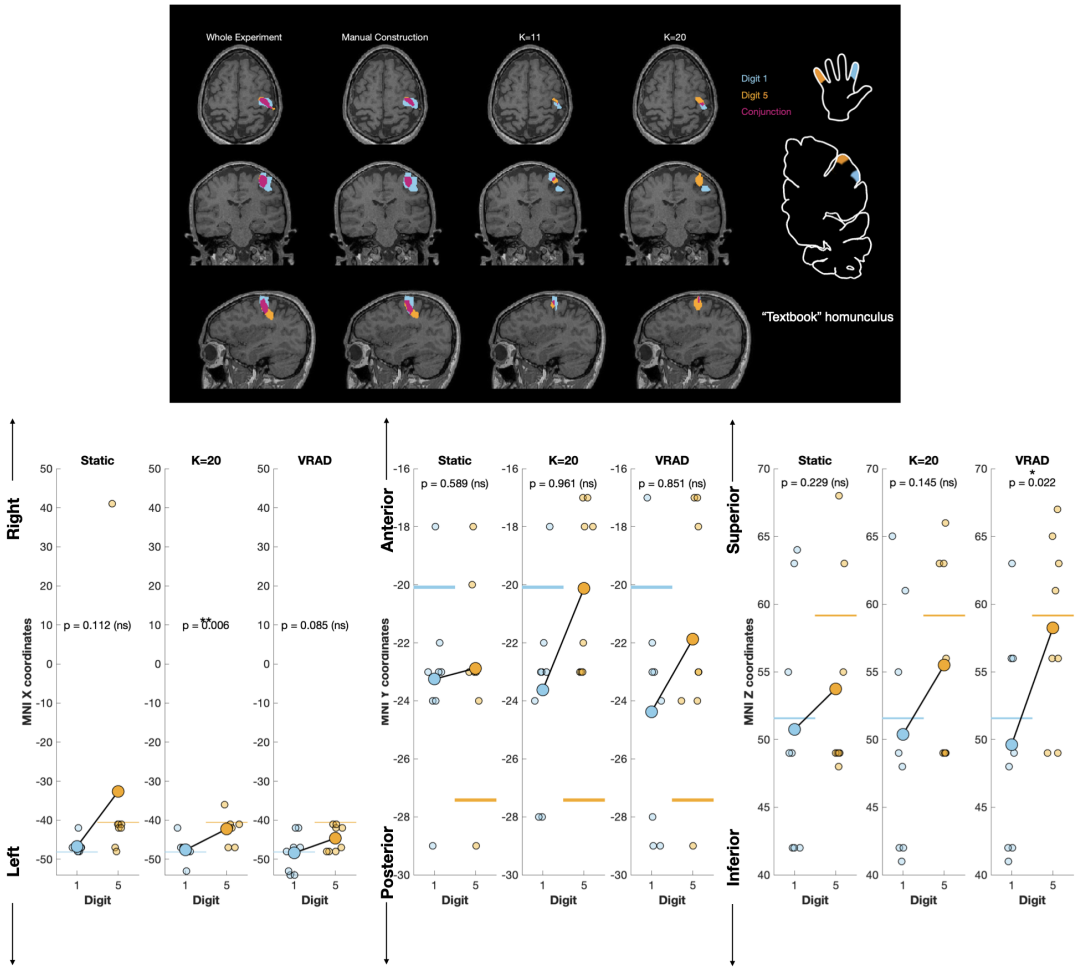


Figure 5.6: Exemplar across session results. Top: conjunction plots of the (multiple comparisons corrected) statistical maps for the different approaches, thresholded at the 90th percentile. Notice the gradual improvement in relation to the anatomical ground truth as more states are introduced, with digit 2 appearing more superior and lateral to digit 1 in the most dynamic case. Bottom: the location of the peak PMBR t-statistics across all sessions for this subject for the static and K=20 state beamformer. The mean location across sessions is circled in black, the centre of mass of a probabilistic fMRI atlas is shown as a horizontal bar. The 20 state beamformer returned statistically significant results in the medial/lateral and anterior/posterior planes. Results were in alignment with fMRI data for the medial/lateral and inferior/superior axes.

Learnt dynamics and spatial modes represent physiological processes

Figure 5.7 shows the inferred spatial modes learnt by VRAD; Figure 5.8 shows the corresponding continuous and epoched α_t mixing coefficients. Note that all epoching was done after inference and the task structure was never (directly) exposed to the

		MNI Coordinate ($\mu \pm \sigma$)		
	Digit	**X**	**Y**	**Z**
Static	**1**	-47 (2)	-23 (3)	51 (9)
	5	-43 (3)	-23 (3)	54 (8)
HMM (K=20)	**1**	-48 (3)	-24 (3)	50 (9)
	5	-42 (4)	-20 (3)	56 (7)
VRAD	**1**	-48 (5)	-24 (4)	50 (8)
	5	-45 (3)	-22 (4)	58 (7)

Table 5.1: Mean peak t-statistic across 8 sessions for highlighted subject in MNI coordinates. The standard deviation is denoted in brackets.

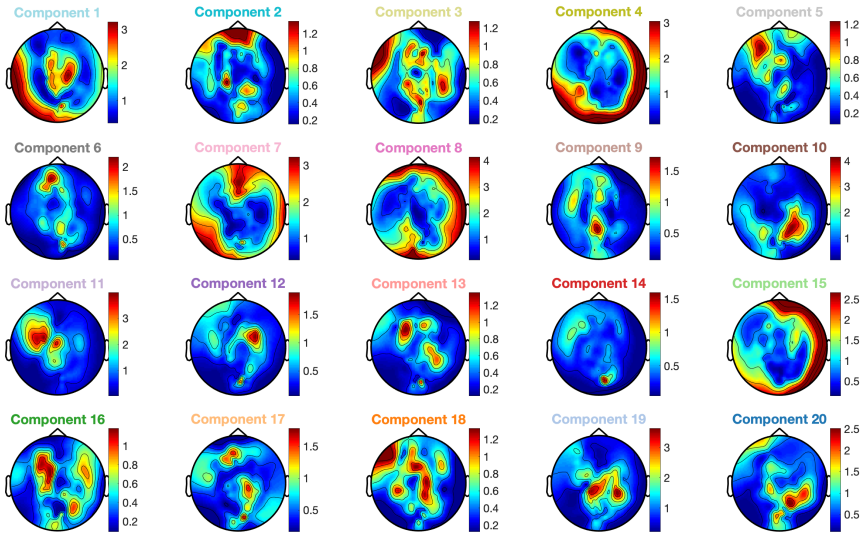


Figure 5.7: Spatial modes inferred by VRAD for one subject's session. Components appear to be a mixture of artefact (1, 4, 7 and 15 - cardiac and 2 - ocular) and physiology (10 and 11), with prominent dipolar patterns visible.

inference architecture. Note also that the timeseries and power maps are shown on their own scale.

The continuous α_t traces suggest that VRAD is able to capture time-varying neural and artefactual contributions to the sensor level covariance; components 1 and 15 exhibit very rhythmic patterns whilst components 10 and 11 activate more sporadically. This notion is supported when looking at the accompanying spatial components, with components 1 and 15 resembling the classic magnetocardiogram (MCG) ringing patterns whilst component 11 peaks over ipsilateral motor sensors with a dipolar pattern. Finally, notice how in the continuous view of the dynamics how more than one component contributes to the final sensor level covariance at any one instance, in stark contrast to the result obtained by a HMM (see Figure 5.11B)).

Further evidence for the idea that VRAD learns realistic depictions of task-induced physiological phenomena is gained once the α_t timeseries is epoched; components 11 and 19 display a "classic" beta rebound (Cheyne, 2013; Kilavik et al., 2013; Pfurtscheller and Da Silva, 1999). Interestingly, state 10 appears to show activation during finger motion.

Arguably, these two different views of the dynamics are most informative when

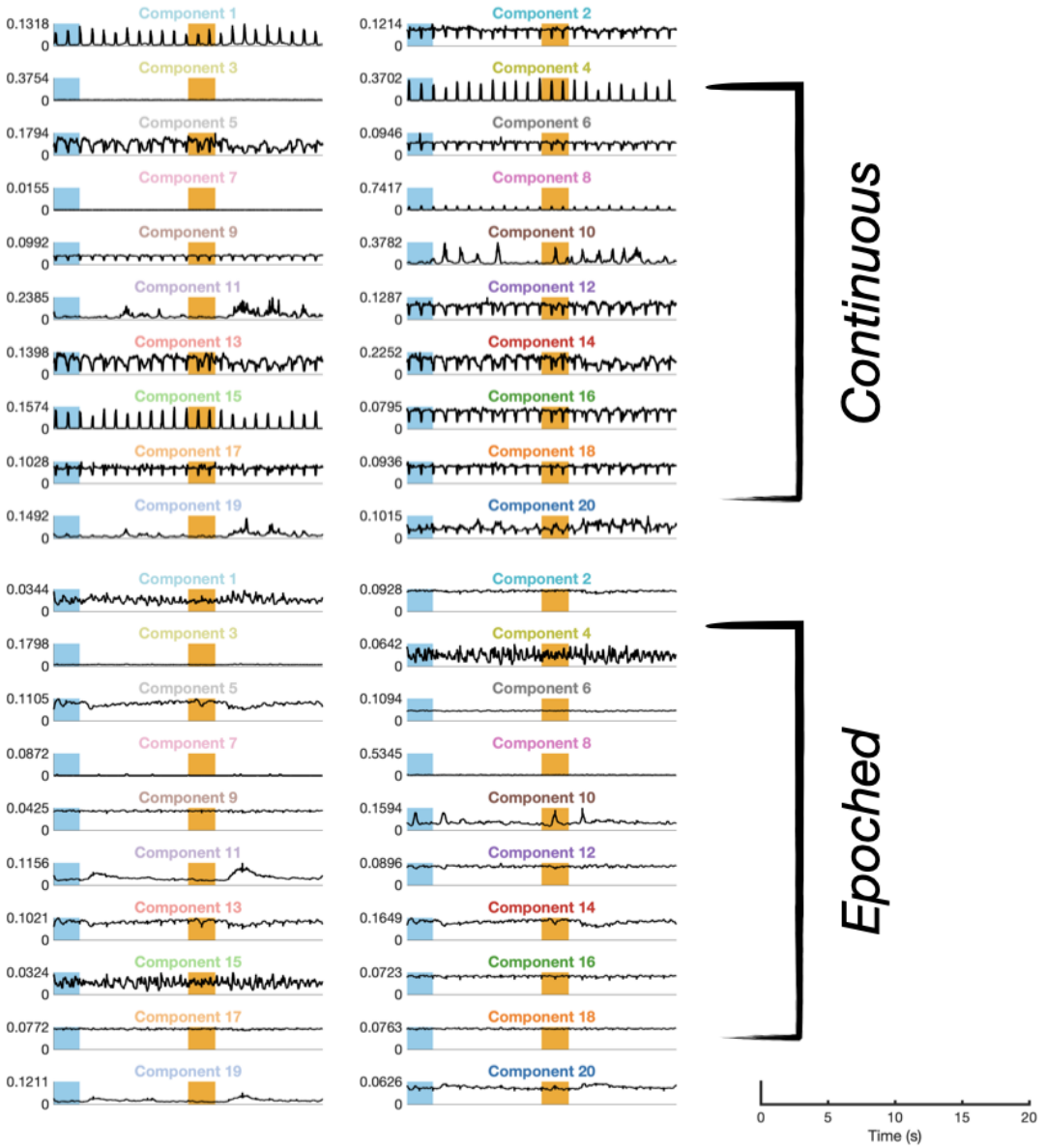


Figure 5.8: Continuous (top) and epoched VRAD mixing coefficients, α_t, with respect to the task structure (bottom) give complementary views of covariance dynamics. Top: rhythmic mixing coefficients are easily identified and suggest non-neural sources of interference (e.g. heartbeat artefact). This view of the data also reveals the relationship between trial dynamics and the continuous task structure, with components 11 and 19 being most prominent after movement cessation. Bottom: epoching the data further helps with classifying which modes are of neuroscientific interest, with a PMBR observable in components 11 and 19. Component 10 (on average) peaks both during and after finger motion. Note that mixing coefficients are shown on their own scale. 20s of data is shown in all plots.

combined together. Figure 5.9 shows the activation of component 11 over all trials as an image with respect to the task structure. Visualising the data implies that state 11 is a “rebound" state. Note however that there is substantial inter-trial-variability exhibited in the recruitment of this component in terms of onset, duration and amplitude. When averaged over trials (white trace), the widely reported beta rebound profile is revealed.

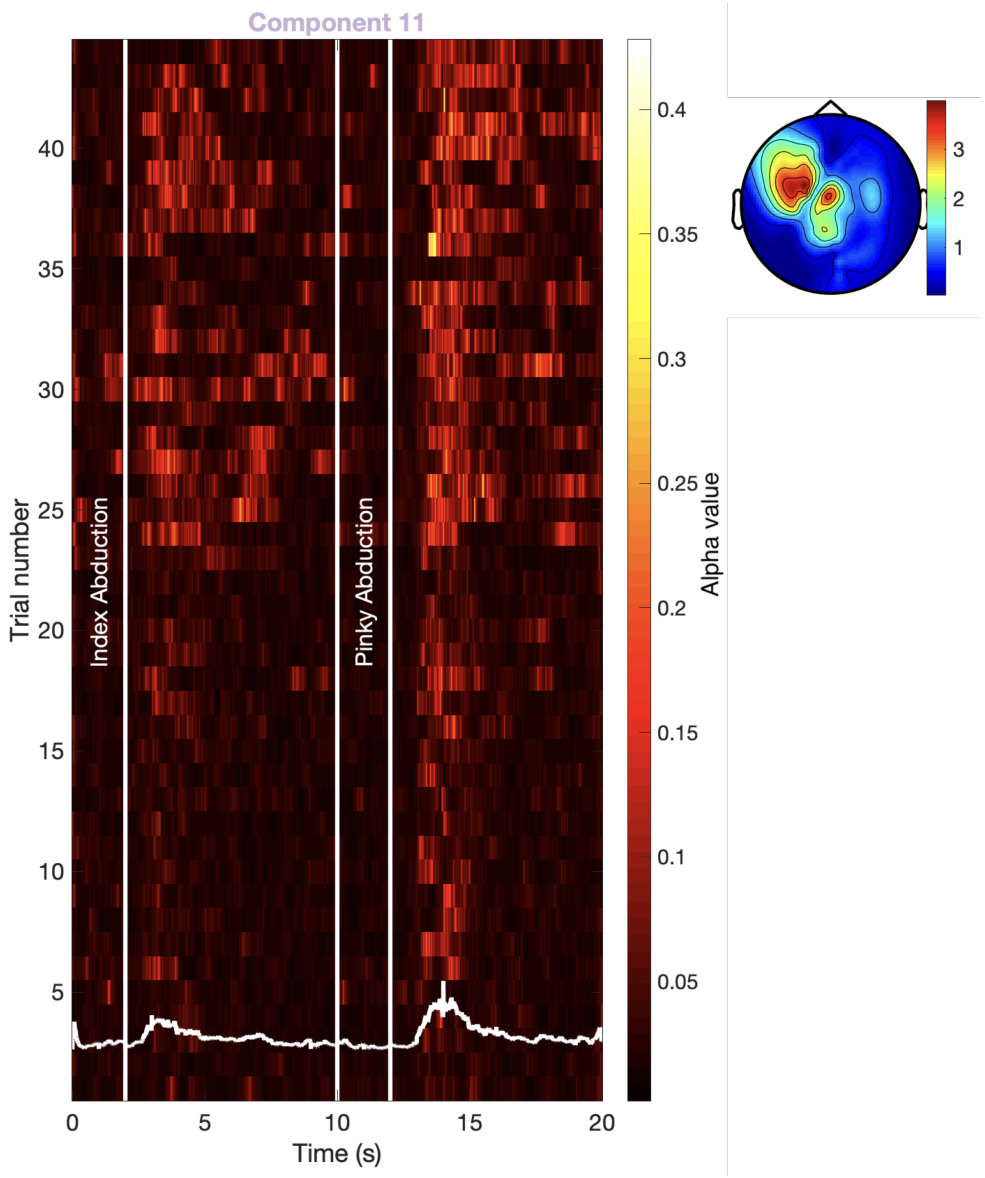


Figure 5.9: VRAD mixing coefficients exhibit substantial inter-trial variability and suggest beta bursts. The PMBR component identified from Figure 5.8 is shown for all trials as an image, with the average of these data plotted as a trace in white. The associated spatial map is also shown.

As VRAD is a non-categorical model, and we place no restrictions on the scale of the inferred spatial components (i.e. the $\mathbf{B}_j$ matrices in Equation (3.4)), looking at the raw α_t traces can obfuscate the reality of what has been inferred. It can be more informative to look at the mixing coefficients after they have been rescaled by the magnitude of their corresponding spatial maps. Figure 5.10A) shows the maximally contributing component to the overall variance as a function of time, over all 44 trials. This was obtained by scaling the α_t coefficients by the trace of the inferred spatial matrices, and taking the argmax for each timepoint. The same colours are used to represent each component as in Figure 5.7.

Rescaling the data this way reveals that component 14 (red) dominates the timeseries with other components (10 and 11, brown and lilac) contributing more around the onset and offset of task. One interpretation of this result is that VRAD finds a "background state" and adds task-specific components to parameterise the melange of covariance when necessary. It should be stressed that this is a consequence of

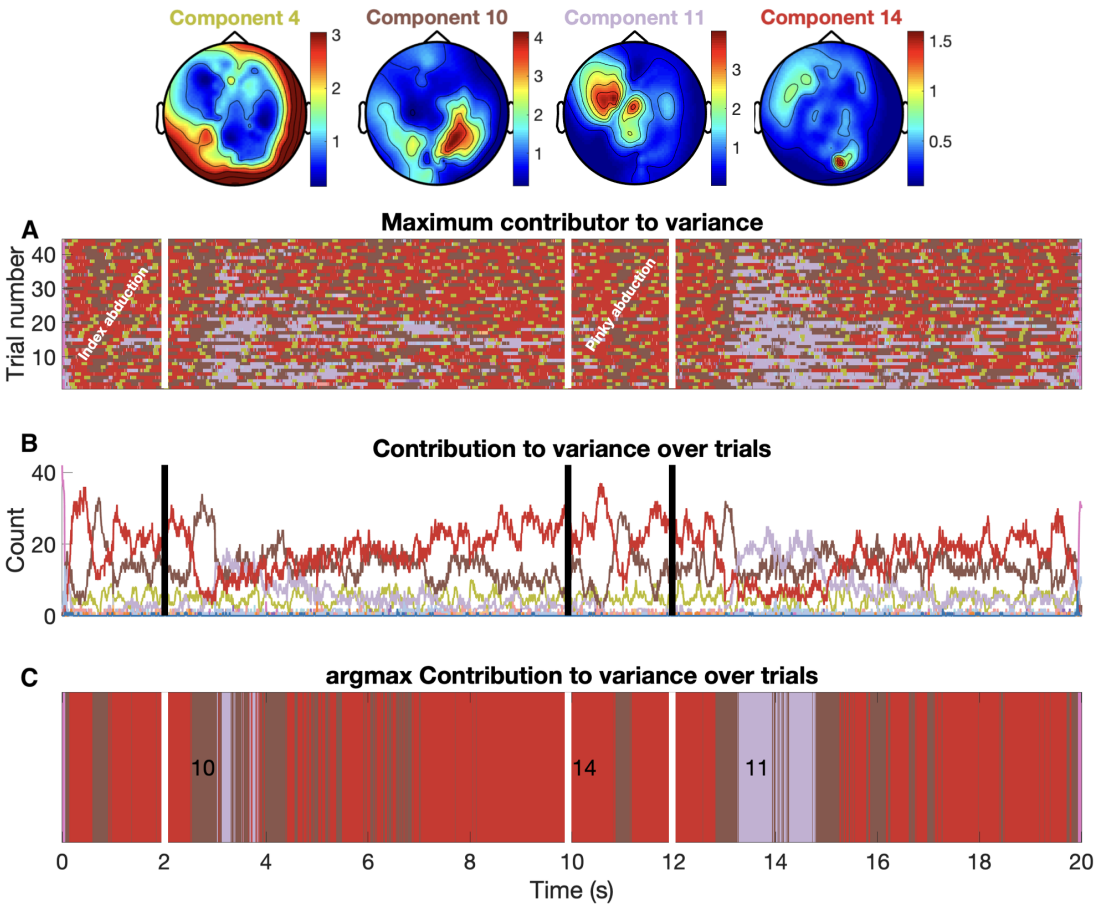


Figure 5.10: Accounting for the scale of VRAD's inferred spatial components reveals the most prominent contributors to the instantaneous power. Panel A shows which component contributes the most to the overall sensor level variance. The column-wise sum of these data are shown in panel B. Component 14 tends to dominate, but is suppressed in the post-movement periods (2s and 12s), at which times components 10 (brown) and 11 (lilac) become more relevant. Intriguingly, component 10 appears on average to precede component 11, as in shown in the medallion plot in panel C (obtained by taking the argmax of those data shown in panel B). Caution must be applied to the interpretation of this result, as the VRAD model allows for both components to contribute simultaneously. This mixing of components is pronounced in the continuous view of the data in Figure 5.8.

the inference and is not setup by design.

Panel B) shows the column-wise sum of the argmax and rescaled data (i.e. those data shown in panel A). A post-movement beta rebound (with respect to motion offset) can again be seen in the lilac trace (component 11). Interestingly, component 10 appears to be recruited during the motion period (0-2 and 10-12s) and immediately precedes the peak of the PMBR state. Panel C shows the argmax of panel B over time and highlights the prevalence of the background component (14, red) and the rebound state.

Comparison with a HMM

Although VRAD and the HMM are fundamentally different models (as detailed extensively in Chapter 3), a qualitative comparison of their results can act as a means to check that the new technique is returning sensible results. Figure 5.11 shows a 20 state HMM result for the same subject and same session as presented in

Figures 5.7-5.10. Panel A) shows the HMM inferred states which activated during the movement-related desynchronisation and the PMBR periods of the task. Left: plots of the inferred (normalised) variance learnt at the sensor level and the associated averaged state time course, obtained by epoching after inference. All averaged traces are shown on the same scale. Notice how states 11, 15 and 19 exhibit a PMBR in the time domain, and peak in their variance maps over left motor cortex in space. Conversely, states 4, 14 and 16 appear to be more associated with motion onset, with their maps peaking more over posterior sensors. The remaining states appear to pick up more subtle components of the data (e.g. ipsilateral activation in state 8) whilst others (e.g. 7 and 12) appear to describe artefactual components such as the cardiac cycle (see the tight distribution of state lifetimes in Figure 5.12). In contrast to the non-stationary approach, when using a $K = 1$ HMM (right, akin to just using the classic covariance Equation) one loses this rich spatiotemporal description of the data.

Figure 5.11B shows the state timecourse for a representative 5 seconds around index finger motion for the $K = 20$ HMM model. Note the rich and varied recruitment of states on the millisecond timescale, with states 14, 15 and 19 being prominent in the rebound period ($\sim$ 2 to 4s) in this example.

To formalise our results, we modelled the mean state time course (timelocked to the experiment). We performed a non-parametric statistical test, correcting for multiple comparisons by means of taking the maximum t-statistic (pooled over all states and all time points) upon each permutation (sign flip) of the design matrix. We ran $N = 1,000$ permutations. Figure 5.11C) (left) shows the statistically significant state activations thresholded at the $p = 0.05$ level; the right hand figure shows the distribution of maximum t values.

Of the 20 inferred states, 8 survive statistical thresholding (over the averaged experimental window), and these appear to correlate with the task structure and what is known about the temporal characteristics of the brain during a motor task. Prominent are states 19, 15, 14 and 4, with the latter two states correlating with movement onset; states 19 and 15 are visited during the typical PMBR time window. Interestingly, on average, there appears to be a sequential nature to the visiting of each state: 14, 19 to 15 for both digits.

Gaining confidence in our results with state activations

Whilst VRAD is a more flexible model, the HMM's categorical nature makes it convenient to easily calculate state-specific statistics. We now exploit this to further validate our findings.

A lot can be learnt from the sensor level topographies, yet a greater intuition is gained when the power map for each state is visualised at the source level. We calculated the root-mean-squared (RMS) power when a state was activated and contrasted this with the RMS power for all other points in time.

The left of Figure 5.12 shows the power maps for the most prominent/statistically significant states. Maps are thresholded at the 50th percentile and centred on the peak absolute value. The top panel shows the baseline power distribution across the whole experiment. We normalised the beamformer weights (static case) when constructing this map to remove bias towards the centre of the head to ease visualisation, but note that this step was not carried out for any of the individual state maps. The bottom panel shows the visually identified "beta desynch states" in blue and "rebound states" in red. We also show a map for one of the artefact states - notice the lack of spatial structure in these maps. State 4 exhibits a clear decrease in power across both hemispheres in the motor cortex region, whilst state 14 recruits both pre and supplementary motor areas. In contrast, states 15 and 19

are localised solely to the left motor region and peak at different voxel locations.

Covariance (and therefore beamformer weights) accuracy is a function of the number of datapoints available to estimate the statistic. It is therefore important to attempt to quantify how reliable the covariance estimate is likely to be for each of the different matrices/basis functions. The HMM naturally provides such a metric in the form of the fractional occupancy (FO); the right of Figure 5.12 shows the FO alongside other state statistics for the 20 state example, with a relatively even distribution of occupancies. For 44 trials, each lasting 20s, this gives $\sim 44s$ of data per covariance matrix state. This has been shown to be an adequate number of data points for beamformer estimation in the past (Brookes et al., 2008).

A similar estimate for the number of contributing time points was also made for the VRAD-powered beamformer. We did this in two different ways. Briefly, first a two component GMM was fitted to a transformed version of the mixing coefficients, with the curve fitted to the smallest values determining when a component was “off". We also tried a more simple method, based off of Otsu thresholding (Otsu, 1979), to binarise the mixing coefficients. Whilst neither of these methods are ideal (see an extensive discussion on this in Chapter 7), and there is a large degree of uncertainty in this statistic, importantly the results in Figure 5.14 show that 17 out of the 20 elements are considered “activated" for more than 30s, meaning that we can have some faith in our covariance estimation. The three elements which did not reach this threshold are likely to be stitching artefacts.

Dimensionality Reduction via K-means clustering

K-means clustering was run to approximate the inferred mixing coefficients learnt by VRAD. This was done to 1) reduce the computational demand of calculating up to T reconstruction weights and 2) regularise the dynamics - by using the centroids of the clusters we are less susceptible to noise in the inferred timeseries. The top of Figure 5.13 shows half a second of the true mixing coefficients and the approximated (discretised) coefficients after running the K-means clustering with $K = 400$ centroids. This led to a percentage error of 9.61 ± 15.85 (mean and standard deviation) across each alpha value and a total sum of squared error of 5.84. We investigated the effects of using different numbers of centroids for the discretisation of the alphas as shown in the bottom panel; 20 logarithmically spaced integers for K were used between 10 and the total number of time points. We chose the smallest number of centroids which resulted in a mean percentage error smaller than 10%.

Discussion

This chapter discusses results specific to this high-resolution dataset; more general discussion points surrounding the technique (e.g. hyperparameter choices) can be found in Chapter 7.

Source Reconstruction Results

Achieving high spatial resolution images in MEG is challenging. This stems primarily from the fact that the MEG inverse problem is strictly-speaking insoluble, requires the introduction of assumptions and is dependent on the solving of the forward problem with high accuracy as a pre-requisite step. Moreover, confounds such as subject motion, coregistration inaccuracies and sub-optimal SNR which are present in traditional MEG experiments compound to add an extra layer of complication to pursuits for spatially optimal statistical maps.

Barratt et al. (2018) overcame the practical hurdle of improving forward modelling by effectively eradicating head motion and coregistration error by using foam head casts. This ameliorated the estimate for the lead fields featured in Equation (2.9). Inspection of this Equation reminds us that the inverse model is also a function of the estimate for the sensor level covariance, as well as the measured data themselves. Inspired by the previously published work, we sought to optimise the modelling of the data covariance by incorporating dynamics into the system. Unlike the previous study however (Barratt et al., 2018), we took a fully unsupervised approach to this temporal modelling. We compared our results with static beamformer and task-informed covariance segregation methods.

We were able to reproduce the results of the original paper (Barratt et al., 2018), showing that simplifying assumptions about the data covariance (that it is stationary) can lead to more erroneous source images than when more than one covariance matrix is used. We did this in the context of a motor mapping experiment and were able to spatially distinguish spontaneous oscillations arising from the cessation of digit motion on the same hand in contralateral primary somatosensory cortex. Importantly, this was only possible in non-stationary approaches, where we showed a statistically significant shift in the location of the peak PMBR T-statistics for the beamformer implementations. We used cross-modality prior information (structural and functional) to add further confidence to our findings. We showed that our dynamically derived images were closer to these historic results. As well as being more accurate (closer to the expected ground truth), we obtained higher resolution images as shown by the reduced overlap in the conjunction maps. All of this was obtained with a mere 44 trials per session.

A very fair critique of the results presented in Figure 5.4B) is that the comparison between the two approaches isn't fair: MNE solutions are famously smooth, whilst MSP is known to give focal source images (Friston et al., 2008; Lopez et al., 2014), even before dynamics are introduced. This is a perfectly valid point - the purpose of this Figure is merely to illustrate the function of the framework developed throughout this thesis. More work needs to be done with fairer comparisons (e.g. MNE vs TAMNE, MSPs vs TAMSPs) if any meaningful conclusions between the two methods (static versus dynamic) are to be drawn.

For computational reasons we opted to reconstruct activity to a 6mm isotropic grid, whilst the original paper used a 2mm one. This of course to some extent limits the spatial precision of source space results. Nonetheless, we recovered a statistically significant shift in the peak PMBR location, and we anticipate that we would have drawn similar (if not the same) conclusions if we used a smaller grid size.

Headcasts

Optimal source reconstruction demands good estimates for the lead field matrix, of which coregistration is a factor. This was the primary motivation for the invention of foam headcasts (Meyer et al., 2017a) which were used in this study, as they facilitate high-precision MEG experiments.

Headcast datasets have minimal coregistration error and suffer negligibly due to head motion. This is particularly pertinent in this study, where the spatial separation of digits was shown to be on the order of 4mm, comparable to the anticipated coregistration error in a conventional MEG experiment (Hillebrand and Barnes, 2011). The controlling of head movement was therefore critical for maximal spatial resolution to be achieved. Despite their advantages however, head casts are not a panacea for conventional MEG, being expensive, uncomfortable and potentially dangerous. Above all, head casts take MEG one step further away from being either a useful clinical tool (patient cohorts with movement disorders are immediately excluded) or a method to study participants in a naturalistic setting.

OPMs

This work demonstrates the ability for a cryogenic based MEG system to spatially disambiguate spontaneous oscillations arising from simple finger motion on digits 1 and 5. However, a great deal of effort was invested for this feat to be achieved, with state-of-the-art forward modelling and source inversion techniques both being employed. Simulation work in the original paper suggested that the maximum obtainable spatial resolution with a cold MEG system was on the order of 2mm, with this resolution degrading when non-white noise was added to the simulation (see Discussion chapter). This means that successfully separating non-phase locked activity for all 5 digits in space is likely to be very hard with a SQUID-based system.

There is currently a palpable sense of optimism about the future role of OPMs within the field of MEG, with a substantial body of numerical work now promising enhanced spatial resolution over cryogenic alternatives. A natural progression of this work would be to repeat the experiment using OPMs, recruiting all 5 digits. However, several hurdles need to be overcome for this to be a success. This includes improvements in noise modelling, accounting for sensor gain error, attaining sufficiently high coregistration accuracy and, if the SNR becomes sufficiently high, dealing with array mismatch.

Accounting for forward model uncertainty with a generative model

Although, as discussed in Chapter 2, one can invest efforts into deriving more and more accurate headmodels, the ground truth forward model is unlikely to ever be recovered. This is true even in the current study where coregistration error is effectively minimised. This can lead to the output SNR of spatial filter results being significantly diminished, a term dubbed “array mismatch" (Sekihara and Nagarajan, 2008).

This effect becomes particularly pronounced at higher SNRs. One way of attempting to deal with this lead field uncertainty is to combine the beamformer with a generative model (Belardinelli et al., 2012). In such instances, different estimates for the lead fields can be used to give a batch of source space estimates, and some model evidence criterion can be used to pick the most optimal forward model. This was the approach pioneered by López et al. (2012), to deal with coregistration error in a normal (non-headcast) experiment. Our TAMSP approach could have facilitated this approach, but we did not explore this option in this thesis.

Beta bursts are identified by the inference

Recently there has been a shift away from the adage that the electrical brain comprises of sustained oscillations towards the notion that power spectra is better understood as consisting of transient “bursting" events (Jones, 2016; van Ede et al., 2018). One of the main health warnings to be taken from this work is that the time-averaged result of spontaneous oscillations should be interpreted with caution, as the effect of averaging many temporally distinct bursts in activity can lead to the erroneous conclusion that the brain is exhibiting sustained rhythms. Most importantly, this bursting perspective of task-responses has shown to be both descriptive and discriminatory in the clinic (Tinkhauser et al., 2017) as well as in the neuroscience laboratory (Feingold et al., 2015; Shin et al., 2017; Seedat et al., 2020).

We were able to observe this phenomena of burst-like events with the new VRAD inference framework, where, on average, the latent variable dynamics resemble a PMBR, yet there was substantial variation of component recruitment in terms of amplitude and duration (Figure 5.9). Coincidentally, the HMM has shown to be one efficient way of detecting such spectral phenomena through use of time-delay

embedding (Quinn et al., 2019). We have therefore helped to further validate the use of unsupervised machine learning techniques in burst detection with a latent variable model. The observation of these bursting events may be missed if one was to adopt the classic one or two window model of covariance for beamformer estimation. Indeed, it appears that accounting for these bursts was important in the construction of higher resolution spatial maps.

Quantifying Covariance Estimation Uncertainty

Brookes et al. (2008) and our simulations in Chapter 4 stress the importance of having adequate data to estimate covariance. We estimated the number of datapoints going into the construction of each covariance matrix using two different methods for the VRAD approach. The results for this estimation is shown in Figure 5.14. Importantly, we were able to show that the lower bound on the number of data points used for covariance estimation was sufficient, a result which also held for the HMM beamformer (Figure 5.12).

Oversimplifying our results

Our post-hoc analyses of the source space data over-simplified the complexity of the results. This is illustrated in Figure 5.6, where we took the peak PMBR t-statistic to assess the quality of the source reconstruction, collapsing data across frequencies (beta band), time (only looking at one statistic derived from the entirety of the PMBR window) and space (inspecting just the peak voxel). Whilst this makes the results more digestible, not only does it overlook potentially interesting effects (especially across different frequency bands, see Dalal et al. (2008); Nowak et al. (2017, 2018)), it also makes our results susceptible to outliers, as demonstrated by the ipsilateral peak activation for the static beamformer case. The conjunction maps used in the top of Figure 5.6 go some way to overcome this limitation, but still requires some user choice (the threshold level of the statistical map).

Cross-modality priors

One of the benefits of the MSP framework is the ability to use structural and/or functional priors, including those derived from different imaging modalities, to drive source reconstruction. This has been successfully employed on numerous occasions (Henson et al., 2011; Wakeman and Henson, 2015). One of the arguments *against* using cross-modality priors is that they *could* bias the inversion technique away from the true source data. Whilst this issue is potentially circumvented by introducing “soft" priors into the model, the original paper showed that the topological organisation of the PMBR is spatially distinct across MEG and fMRI, and so we opted not to use the fMRI derived z-stats in this work.

Conclusion

We have shown that MEG can recover elements of the human homunculus from spontaneous oscillations arising from post-movement-cessation with the help of a temporally adaptive source reconstruction pipeline. We demonstrated this in two different settings - in a generative based model approach (TAMSP) - and in the well-loved spatial filter regime (TABFER). Comparisons and improvements were shown over static counterparts. Our spatial filter results show that accounting for time-varying neural phenomena (e.g. beta bursts) leads to the construction of more accurate source space images. Remarkably, these results were obtained with a mere 44 trials per session.

Acknowledgements

We would like to thank the Nottingham MEG lab for providing us with their data. We would also like to give a heart-felt thank you to Jimmy Bonaiuto for all of his help with FreeSurfer right in the middle of a pandemic.

The computational aspects of this research were supported by the Wellcome Trust Core Award Grant Number 203141/Z/16/Z and the NIHR Oxford BRC. The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health.

Appendix 5.1: TAMSP Pre-Processing

FreeSurfer's recon-all command was run on the individual's MRI to extract the grey and white matter surfaces (Fischl, 2012). Hemispheres were combined and the resulting mesh was downsampled, resulting in roughly 60,000 vertices left in the source space.

We used SPM to preprocess the data. After the lead field was calculated, MEG signals were projected to a smaller space by performing a singular value decomposition on the lead field matrix: $\mathbf{USV}^T = \mathbf{HH}^T$. This gave the spatial transform $\mathbf{A} = \mathbf{U}^T$, $\mathbf{Y} \rightarrow \mathbf{AY}$. Spatial priors were obtained by splitting the four motor regions in the Desikan-Killiany atlas into 25 smaller constituent parts (Desikan et al., 2006). A nearest neighbour interpolation of the atlas onto the subject's mesh defined each of the patches, β_i. Note that unlike in the default SPM settings we just used a diagonal covariance matrix for each prior. Finally, we combined all remaining non-motor prior patches into one "rest of the brain" source space prior (effectively amounting to a MNE solution) to soak up the rest of the neural variance. These priors were projected to the sensor level through the reduced lead field, $\mathbf{AH}$. We used a single IID noise prior to model the noise. These 102 priors, along with the reduced data, were fed into VRAD.

Unlike in the TABFER mode of the inference, we did not restrict the mixing coefficients to sum to 1, but did ensure that they were positive with a "softplus" constraint. Also note that no initialisation was carried out in this mode of TASER.

Upon convergence, we reconstructed data to the motor patches using Equations (4.2) and (2.10). We did not do any temporal downsampling to the coefficients in the example presented here. T-statistical images were made as per the beamformer implementations. No multiple comparisons correction was carried out on this example. For a static comparison, we also reconstructed activity using SPM's MNE implementation.

Appendix 5.2: The importance of covariance inversion regularisation

Figure 5.15 shows the same statistical image as in Figure 5.5, but this time without any eigenvalue truncation applied in the matrix inversion step, $\mathbf{C}_Y^{-1}$. Notice the noise induced into the images. This is easily accounted for with some basic linear algebra tricks. Briefly, an eigenvalue decomposition is carried out on each covariance matrix, $\mathbf{U\Lambda U}^{-1} = \mathbf{C}$. Using the fact that $\mathbf{C}^{-1} = \mathbf{U\Lambda}^{-1}\mathbf{U}^{-1}$, an approximation can be made to calculate the inverse. The first κ eigenvalues and eigenvectors are kept, once any redundant or troublesome eigenvalues (those which are equal or close to zero) are removed: $\mathbf{C}^{-1} = \mathbf{U}_k\mathbf{\Lambda}_k^{-1}\mathbf{U}_k^{-1}$.

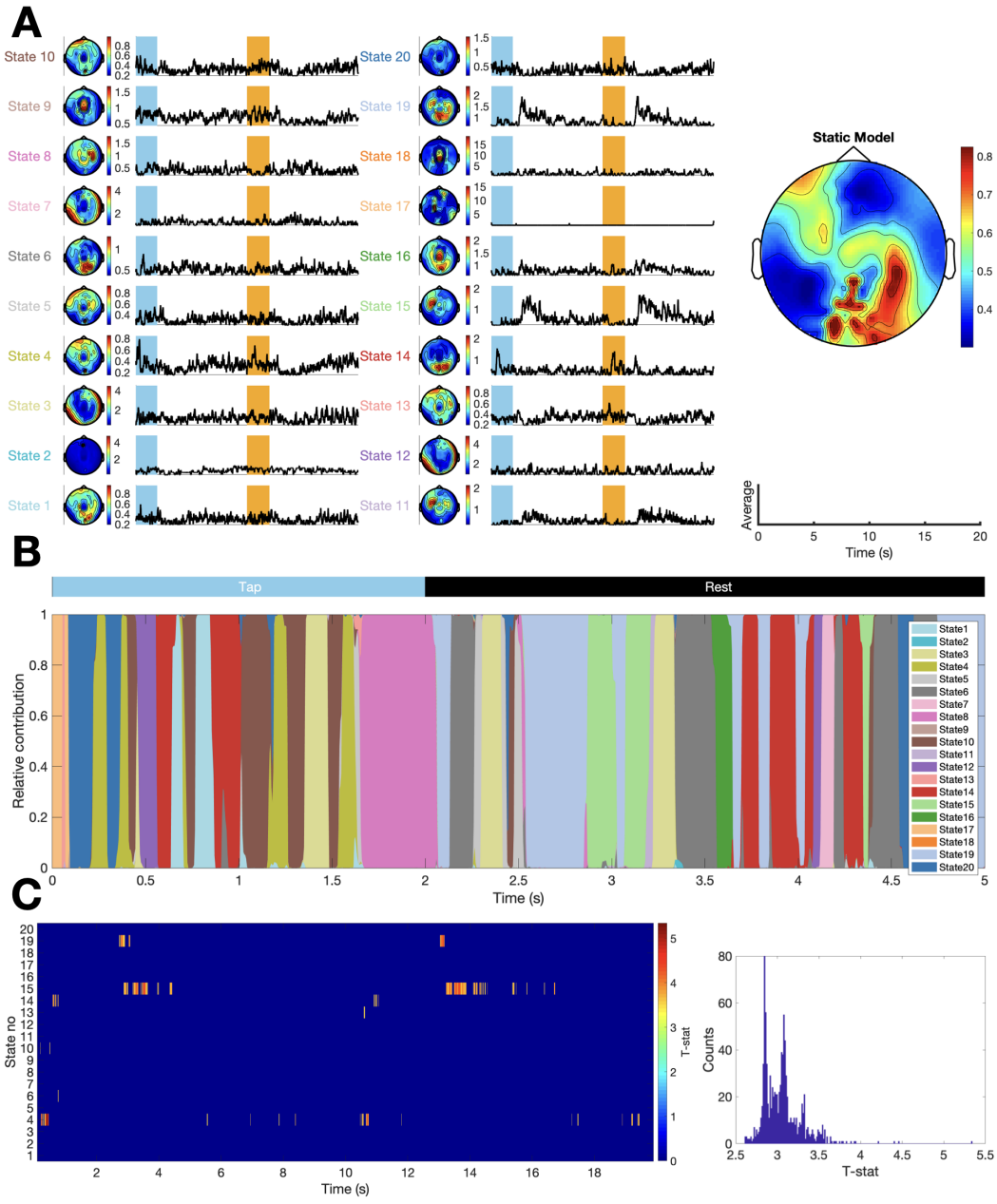


Figure 5.11: Results for the $K = 20$ HMM model. Panel A) shows the inferred spatial maps and their associated epoched gamma timecourses. A $K = 1$ HMM model is shown for comparison. Note, like in VRAD, how the HMM (once epoched) recovers a trace which resembles a PMBR, and that dipolar topographies peak over motor cortex. Panel B) shows the rich recruitment of different states at the start of a finger tap trial, with states 15 and 19 becoming pronounced post-motion in this particular example. Panel C) shows the results of statistical testing on the evoked gammas: blue colours indicate non-significant states and timepoints. States 4, 14, 15 and 19 survive statistical thresholding and correction for multiple comparisons. The right hand Figure shows the null distribution of t-statistics calculated by permutation of the design matrix.

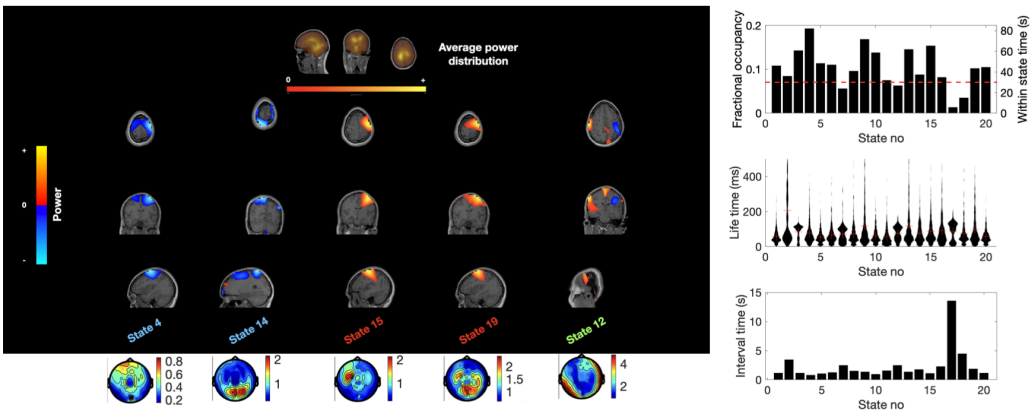


Figure 5.12: State timings give further confidence in dynamic source reconstruction. Left: HMM state-specific power maps, relative to the average background power (top) across the entire experiment. Briefly, data were pooled when states were activated before the RMS power was calculated. Maps are thresholded at the 50th percentile. States 4 and 14 show a decrease in power over motor cortex, with state 14 exhibiting more refined and lateral desynchronisation profile on contralateral motor cortex (in comparison to state 4). Conversely, states 15 and 19 show an increase in power relative to baseline. We did two things to control for the fact that these maps might arise due to chance. First, we visualised the average power distribution (after normalising the lead fields) over the duration of the experiment (top). This was sufficiently isotropic with no clear peaks over motor areas. Second, we used the same method to calculate a power map for an artefactual state; the map shown for state 12 (a heartbeat state) has no discernible spatial structure in brain space, giving evidence to suggest that the temporal structure provided by the HMM state time course drives the spatial structure in the presented images. Right: HMM state statistics quantify beamformer weights accuracy. Top: Fractional occupancy/time within each state for this session. 16 out of 20 states have at least 30s worth of visits, a number deemed more than sufficient for accurate beamformer estimation. Middle: violin plot of the distribution of state lifetimes. The mean is shown in red crosses. Physiological states all have mean lifetimes around $\sim 50ms$. Bottom: state interval times. The artefactual states are easily identified.

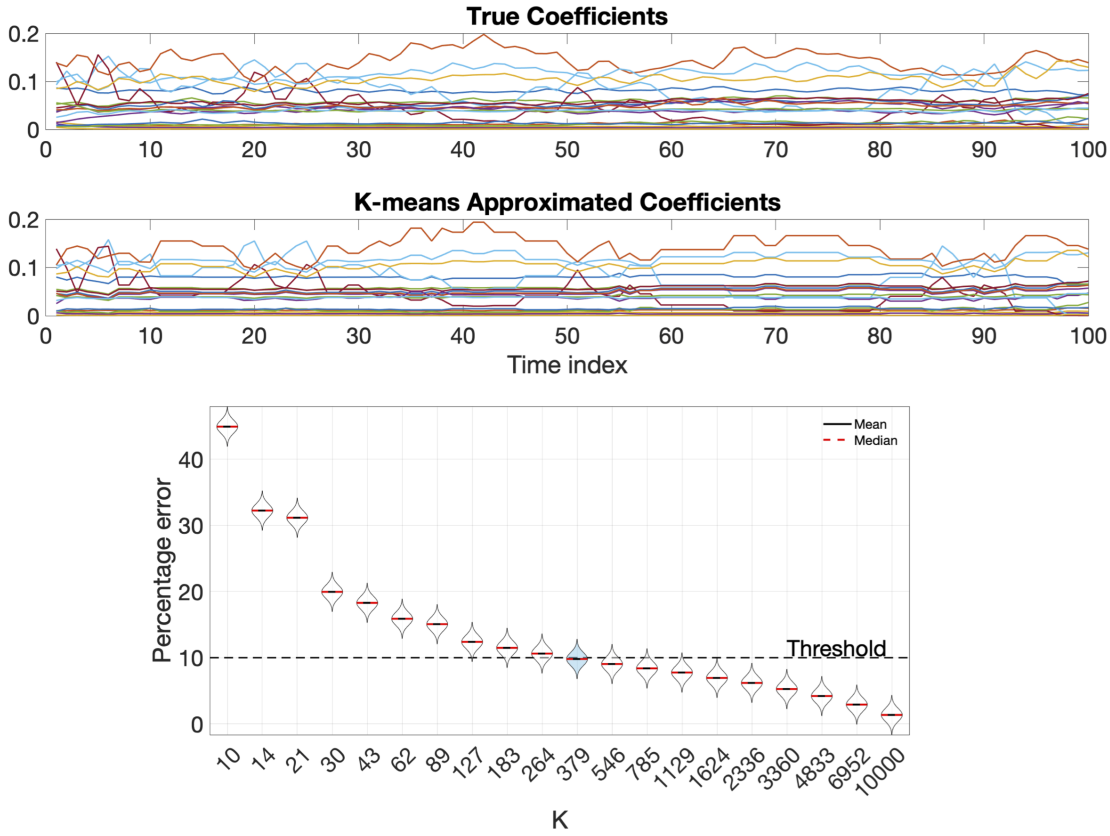


Figure 5.13: Discretising inferred α_t values with K-means clustering. This reduces the dimensionality of the data and regularises the learnt solution. Top panel: inferred coefficients are shown alongside their discretised counterpart for half a second of data. Notice the close resemblance and small error. Bottom panel: investigating how the percentage error varies with the number of clusters. The distribution of absolute percentage error is shown for each cluster number. We choose the smallest number of clusters which gives a mean percentage error less than 10% for all of the $\alpha_{j,t}$ values (highlighted in blue).

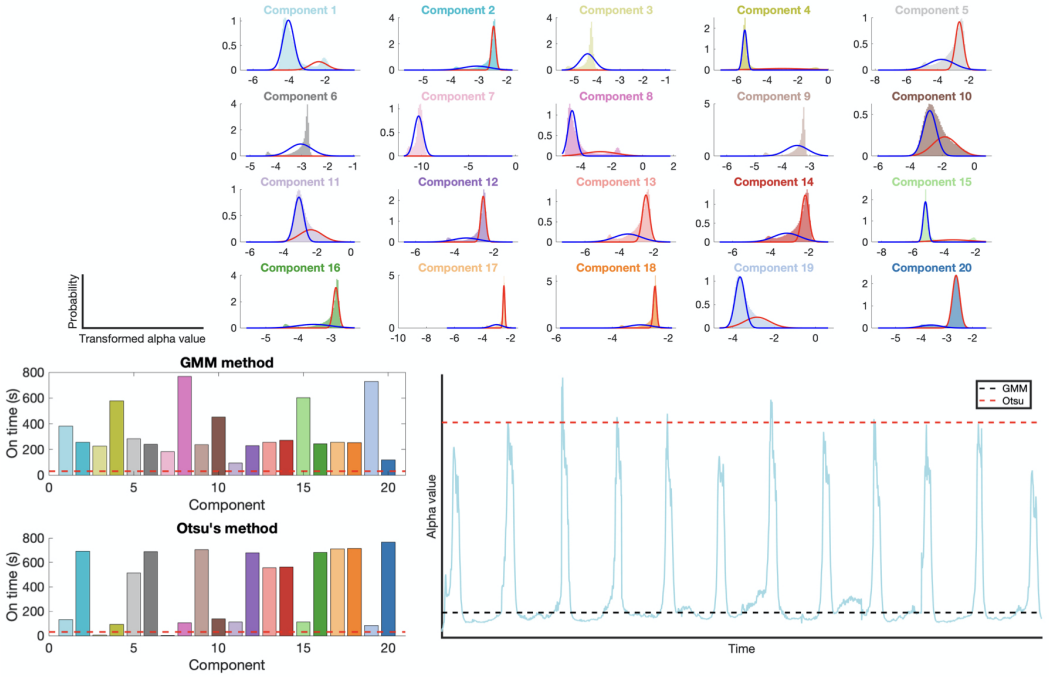


Figure 5.14: Attempting to derive mixing coefficient statistics with a Gaussian Mixture Model (GMM). Top panel: a two component GMM is fitted to the processed (outliers removed and transformed via reciprocation) mixing coefficients. Red and blue curves show the distributions as to when the component was considered off and on, respectively. Taking the argmax of these probabilities for each time point, we are able to approximate the amount of time that each component was on (bottom left panel), and therefore quantify our certainty for the estimation of each covariance element. The bottom right panel shows a 2 second segment of data for component 1, with it's threshold value overlaid in black.

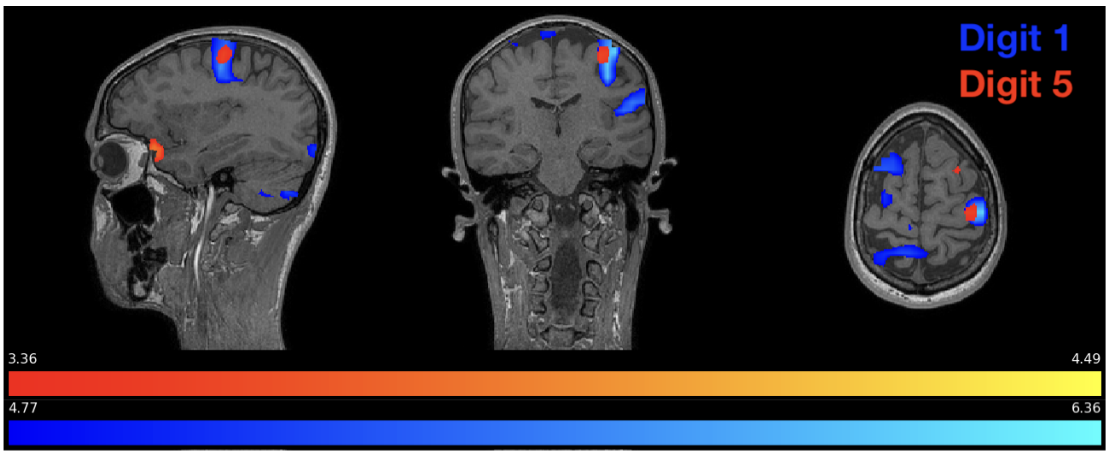


Figure 5.15: The consequences of failing to remove eigenvectors with very small eigenvalues in the matrix inversion step. Trying to invert a close to singular covariance matrix leads to noisy source space images.



6

A more challenging dataset: reconstructing deep cortical activity

Introduction

Having established that a TASER philosophy produces sensible results in Chapter 5, we now seek to test the approach on a more generic and practically challenging MEG dataset.

Generally speaking, cryogenic MEG experiments do not make use of foam headcasts, rendering forward models imperfect due to coregistration errors. Moreover, the neuroscientist does not limit themselves to asking scientific questions exclusively concerning brain regions with good spatial coverage.

Moving away from the idyllic setup of low/no coregistration error and high SNR data, in this chapter we assess how TASER performs on a more “standard" dataset. We use the publicly available Wakeman and Henson Face Perception dataset and run the beamformer flavour of TASER on it. Qualitative and quantitative comparisons are made between these approaches and conventional static source reconstruction techniques.

Background

Face recognition and perception are integral to social cognition and are very well established domains of experimental psychology. This has been studied extensively using modern neuroimaging technologies (see Vuilleumier and Pourtois (2007) for a good review), including MEG (Liu et al., 2002; Kloth et al., 2006; Streit et al., 1999; Perry and Singh, 2014). A classic finding of electrophysiological face studies (in healthy subjects) is a strong evoked response component 170ms post stimulus onset, termed the M170 (or N170 in EEG, (Bentin et al., 1996)). This has been shown to be significantly larger than responses to non-face stimuli at the same latency (Liu et al., 2002, 2000) and is believed to be the brain’s response to the encoding of faces (Haxby et al., 1996). fMRI studies and M/EEG source imaging have identified the fusiform gyrus, lateral occipital cortex and superior temporal sulcus as the key brain areas involved in face processing.

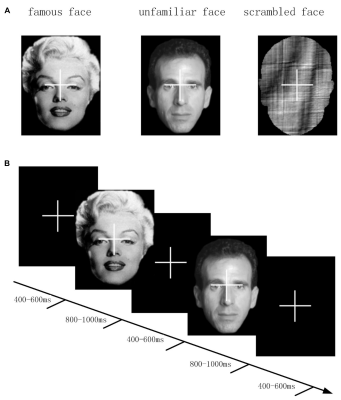


Figure 6.1: Exemplar stimuli presented in the Wakeman and Henson dataset. Figure taken from Wang et al. (2019).

Dataset Synopsis

Here we analyse 16 subject's magnetometer data from a publicly available Elekta Neuromag dataset (Wakeman and Henson, 2015), in which they were presented pictures of famous, unfamiliar or scrambled faces. Participants were scanned for 6 separate sessions, giving ∼ 800 trials and totalling approximately one hour's worth of data. A button press was included in the paradigm but is not analysed here.

As detailed extensively in Chapter 2, the M/EEG inverse problem is inherently ill-posed. Some researchers believe that the insolubility of this problem is diminished by introducing information originating from different imaging modalities. Researchers acquired the data analysed here with this idea in mind; recording fMRI and EEG data to complement the MEG source inversion efforts. Whilst we do not make use of such data in this chapter, previously published work has shown that using cross-modality information and/or constraining solutions at the group level can improve accuracy (Wakeman and Henson, 2015; Henson et al., 2010, 2011; Sato et al., 2004).

Just like in Chapter 5, we are again here armed with a strong hypothesis as to which brain regions are likely to be evoked in response to the presented paradigm, allowing us to make some form of assessment of TASER's performance.

Previous analyses of this dataset

Fittingly, this dataset is often the go-to option for new source inversion tools. Its analysis features heavily in the SPM manual, where incorporation of endogenous (fMRI) priors are used to drive the source reconstruction within a variational Bayesian framework. Henson et al. (2011) stressed the importance of integrating cross-modality and cross-subject information to regularise and improve the inverse solution. Away from source reconstruction, these data were recently analysed by Quinn et al. (2018) in the source space, where they used a HMM to study the brain's response to task in an unsupervised fashion. A key finding from this work was that the state's inferred by the HMM correlate with the brain's response to face processing on the millisecond timescale. We marry these two previous efforts together here, running our new unsupervised learning tool (VRAD) at the sensor level in the hope of identifying time-varying elements of the data covariance, allowing us to construct timepoint specific reconstruction weights.

Methods

Full details of the data acquisition and experimental design can be found in (Wakeman and Henson, 2015). Relevant details are provided below. We analysed Version 1.04 of the dataset, downloaded from the openneuro repository (https://openneuro.org/datasets/ds000117/versions/1.0.4).

Ethical Consent

All research was conducted in accordance with the Declaration of Helsinki. Participants were members of the MRC Cognition & Brain Sciences Unit participant panel in Cambridge who responded to an invitation to participate. 19 people completed both MEG and MRI visits, of which 8 were female, and 11 were male, with an age range 23-37 years. The version of the dataset analysed here contained 16 of these original subjects. The study was approved by Cambridge University Psychological Ethics Committee. Written informed consent was obtained from each participant.

Experimental Paradigm

The face stimuli comprised two sets of 300 greyscale photographs, half from famous, “familiar", people and half from non-famous, “unfamiliar", people (unknown to participants). A third category of stimulus, “scrambled" faces were generated by taking the 2D-Fourier transform of the faces, permuting the phase information, and then inverse-transforming back into the image space. The shape and size of the scrambled faces were matched to the famous and unfamous images. Half of the faces were male, half female. The famous faces were selected in order to be recognized by the majority of British adults. Non-famous faces were approximately matched to famous faces in terms of their sex and age (by visual judgment). All photos were matched and cropped to show only the face.

Stimuli were projected onto a screen approximately 1.3m in front of the participant; photographs were presented against a black background, with a white fixation cross in the centre. Trials started with the appearance of a fixation cross for a jittered duration between 400 and 600 ms, after which the face or scrambled face was superimposed for a random duration between 800 and 1,000 ms. The interstimulus interval comprised a central white circle for 1,700 ms. Participants were told to fixate centrally throughout the experiment, with the change from central circle to central cross helping to prepare the participant for each stimulus. They were also instructed to try to not blink during the cross-hair or stimulus (but to blink freely during the circle). The experimental paradigm is shown in Figure 6.1.

MEG Acquisition

MEG data were acquired in a light magnetically shielded room using an Elekta Neuromag Vectorview 306 system (Helsinki, FI). A 3D digitizer (Fastrak Polhemus Inc., Colchester, VA, USA) was used to record the locations of HPI coils relative to three anatomical fiducials to facilitate coregistration. Stimuli were presented in six, 7.5 min runs. Data were acquired at 1.1kHz with a lowpass filter at 350 Hz and no high-pass filter.

Structural Data

MRI data were acquired using a Siemens 3T TIM TRIO (Siemens, Erlangen, Germany). A standard 1 mm isotropic T1-weighted image was acquired using an MPRAGE sequence (TR 2,250 ms, TE 2.98 ms, TI 900 ms, 190 Hz/pixel; flip angle 9°).

Preprocessing

We analysed the data which had undergone MaxFilter (version 2.2) processing provided in the openneuro repository. Relevant to the source reconstruction work carried out here, head motion was compensated for every 10 ms within each run and data were aligned across runs to match the head position at the start of the fourth run (the range across participants of movement between runs was 0.2 to 7.9 mm, median = 1.9 mm).

Functional data were pre-processed using SPM12. Magnetometer data (102 channels) were downsampled to 250Hz before being band-pass filtered between 1 and 48Hz using a 5th order Butterworth filter. To remove the risk of biasing our results away from the best possible beamformer image, in the static reconstruction implementation (see below), we epoched the data from -1 to +2s pre/post stimulus. HMM and VRAD analyses were run on the continuous timeseries before being epoched after inference. Bad segments were not removed, nor were artefactual components caused by the cardiac cycle or eye movements.

Structural MRI data were preprocessed using the FieldTrip toolbox. To facilitate the construction of an accurate headmodel, the (de-faced) anatomical MRI for each subject was segmented into grey matter, white matter and CSF using SPM8. As the MR data were anonymised prior to distribution, we did not use the digitised head points provided with the data which included points around the nose, but instead just used the known fiducial coordinates to linearly align the MEG and structural image. A Nolte Single Shell (Nolte, 2003) volume conduction headmodel was used for each subject and a 8mm isotropic grid (bounded by the inner skull) defined the discretised source space. This grid was non-linearly aligned from MNI to the individual space, such that there was a one-to-one correspondence between voxels over subjects. The lead field for three elemental currents was calculated at each of these locations; the rank of the resulting vector at each voxel was reduced to two to eliminate the weakly contributing radial component (Ahlfors et al., 2010; Hämäläinen et al., 1993).

Our complete analysis pipeline was very similar to the one presented in Figure 5.2 (with the addition of the SSS MaxFilter step performed prior to our investigation). These steps importantly included the data rescaling and dimensionality reduction operation. We opted to focus our investigation into the magnetometer data. The motivations for this were two-fold: 1) by using just a single modality, we are able to side-step any issues regarding the rescaling and fusing of data and 2) to try and probe deeper brain structures, a task more suited to this sensor type over gradiometers.

Source Reconstruction

We use the same notation as elsewhere in this thesis to define our terms: $\mathbf{Y}$ is the sensor level data, $\mathbf{C}_Y$ is it's associated covariance matrix and $\mathbf{h}$ is a column of the lead field matrix.

We opted to use the unit-noise gain constraint in this work to remove bias from towards the centre of the head. This amounts to modifying the classic LCMV constraint (Equation (C.1)) such that the weights satisfy

$$\mathbf{w}_{UNG,i}^T \mathbf{h}_i > 0 \quad \text{and} \quad \mathbf{w}_{UNG,i}^T \mathbf{w}_{UNG,i} = 1. \tag{6.1}$$

Under these conditions, the weights are given by

$$\mathbf{w}_{UNG,i} = \frac{\mathbf{C}_Y^{-1}\mathbf{h}}{\sqrt{\mathbf{h}^T\mathbf{C}\mathbf{h}}}, \tag{6.2}$$

which is equivalent to

$$\mathbf{w}_{UNG,i} = \frac{\mathbf{w}_{LCMV,i}}{||\mathbf{w}_{LCMV,i}||}, \tag{6.3}$$

the classic LCMV Equation divided by its norm. This returns the neural activity index (NAI) (Van Veen and Buckley, 1988; Vrba and Robinson, 2000) and is a unitless quantity. This approach was motivated by work suggesting that the unit-noise gain is a good weights normalisation approach when data are likely to be noisy (as is often the case with magnetometer data) (Neugebauer et al., 2017).

We used similar settings to Chapter 5 for the different reconstruction approaches. The regularisation parameter was set to 5% of the mean eigenvalue of each covariance matrix. We set the eigenvalue truncation parameter, κ, to 48. Eigenvalues of the three differently derived covariance matrices are shown in Figure 6.9. Data were analysed on a session-by-session basis (96 separate VRAD and HMM models were run).

1. **“Static" Beamformer**
   All data across all trials (roughly 147 per session) were used to define the covariance matrix, giving $N_w = 1$ weights matrix. Data were de-meaned, and the mean of the pre-stimulus baseline period (-1 to 0 seconds) was subtracted from the data on a channel-by-channel basis for baseline correction.

2. **HMM Beamformer**
   We ran a HMM with a (zero-mean) Gaussian observation model on the pre-processed sensor level data. Prior to inference, we reduced the dimensionality of the timeseries by keeping the top 50 principal components. This typically resulted in 97% of the variance being explained. Importantly, we also normalised the data such that they had zero mean and unit variance on a channel-wise basis. We inferred a model with $K = 20$ states. Upon convergence, the learnt covariances were then projected back out to the native $N_{channels} \times N_{channels}$ space and rescaled to their original variance using the Hadamard product. These covariance matrices defined $N_w = 20$ weights matrices. We then used the Viterbi path from the learnt HMM state time course to determine which set of reconstruction weights to use at what points in time.

3. **VRAD Beamformer**
   We also used the novel inference framework to estimate the instantaneous covariance across the whole experiment (using the same data dimensionality reduction steps detailed above for the HMM-powered dynamic inference). As in Chapter 5, we used K-means clustering to learn 400 centroids of the alpha coefficients, thus giving $N_w = 400$ source reconstruction weights.

   We inferred 20 basis functions. The same settings reported in Chapter 5 for the deep learning model were used.

The absolute value of the source currents were taken after data were projected into the source space. This was to account for the sign ambiguity present in spatial filter results. After construction, statistical images (detailed below) were non-linearly aligned back to MNI space using SPM8.

Statistical Analysis

To limit the scope of our investigation, we only explored the |faces| - |scrambled faces| event related fields (ERF) contrast. This is because the other contrasts can be hard to interpret, and we anticipated the largest signal change in this condition.

First, ERF data from all 6 sessions were constructed and concatenated to give a grand average ERF for each subject and each condition. ERFs were baseline corrected using the 200ms immediately preceding stimulus onset. These data were

then used to construct a two-sided dependent samples T-test between the faces and scrambled faces condition at a latency of 168ms. Statistics were corrected for multiple comparisons over space using cluster-based randomisation ($N_{permutations} = 1000$) as implemented in FieldTrip (Maris and Oostenveld, 2007).

To aid visualisation and interpretation of the source space images, we downloaded a meta-analysis result of studies containing the phrase “Fusiform Face Area" from Neurosynth (Yarkoni et al., 2011). These are presented on the template anatomicals below in the blue/green colourmap, where green indicates a more commonly reported result in the meta-analysis.

Results

VRAD gives a different image of dynamic covariance

Figure 6.2 shows one second of MEG data post stimulus onset. The absolute magnetic field is shown as a topography plot for $t = 116ms$, marked with a black vertical line. A HMM and VRAD fit to these data are also shown; the middle panel shows the probabilistic state time-course derived by the HMM and the bottom panel shows the VRAD mixing coefficients. Note the greater mixing of components in the VRAD case. Also note how the instantaneous VRAD variance more closely resembles the MEG sensor data than the HMM variance. This is obtained when several of the VRAD elements (different coloured lines) add as a weighted sum, whilst the HMM represents the variance as a single state (yellow line) for this time point. The yellow state in the HMM emphasises power at the back of the MEG dewar, but VRAD’s depiction of the data is more similar to the dipole(s) observed in the MEG data at this instance in time.

Source Reconstruction Improves the SNR of Evoked Responses

As expected, all spatial filters dramatically improved the SNR of the sensor level data. This is shown in Figure 6.3 where the grand average (over all subjects and all sessions) of the faces minus scrambled faces evoked response is shown for half a second for the sensor level data alongside the static and HMM implementations of the beamformer. Notice the much clearer response in the static and HMM beamformer plots than in the sensor level data in the top panel at around 168ms (marked by a black vertical line).

This result is also visible when the ERFs for the peak sensor/voxel are plotted, as shown in Figures 6.5 and 6.6. Whilst there is a clear delineation at around 170ms between the two conditions at the sensor level, the difference between the traces is much clearer at the peak voxel in the FFA (identified from the HMM beamformer source space map).

Source Reconstruction Elucidates the Role of the Fusiform Gyrus in Face Processing

The stationary beamformer did return a positive activation in the right FFA. This suggests that the FFA has a stronger response to faces over scrambled/non-face images. However, the statistical map did peak outside of the head (shown by the pink arrow in Figure 6.3). In contrast, the HMM beamformer peaked solely in the right FFA, with results being more focal after statistical thresholding. This is shown in Figure 6.3 where the raw (not multiple comparisons corrected) maps are shown on the same colourscale.

The bottom of Figure 6.4 shows the t-statistical map thresholded at the 75th percentile

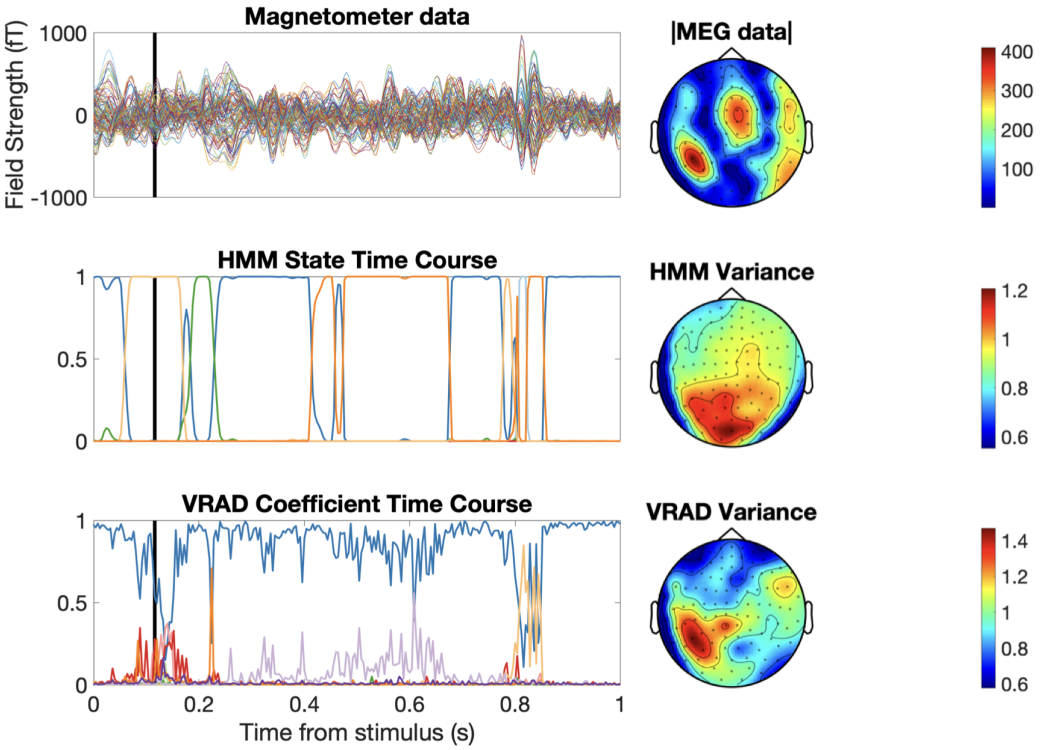


Figure 6.2: VRAD infers plausible latent descriptions of MEG data. Left, top: data from the 102 magnetometers for one trial post stimulus onset is shown. Left, middle: inferred HMM state time course. Different colours denote different states. Left, bottom: inferred VRAD coefficients, α_t. There is no colour correspondence between the traces shown in both timeseries. As expected, the HMM learns a rather binarised description of the data, whilst VRAD facilitates a mixture of covariance components. The black vertical lines on the plots are shown at 116ms post stimulus onset. The corresponding sensor topography and learnt (instantaneous) variances are shown in the right column. The VRAD result is derived from the weighted sum of the covariance elements (i.e. the coefficients shown in the timeseries plot). Note also that the absolute values of the MEG data are shown to ease the visual comparison between the strictly positive variance maps.

for t=168ms in MNI space for the VRAD-powered beamformer. The top plot shows the clusters which survived the correction for multiple comparisons on an individual's MRI. Like in the HMM beamformer result, the map peaks in the right FFA, but a cluster is also found in the left lingual gyrus.

TABFER Recovers the Classic M170

An important validation of the TASER method is to check that it can faithfully reconstruct neural waveforms without distorting the signal. We expect to see a strong evoked response component 170ms post stimulus onset in both the source and sensor space as a response to the presentation of a face. (Bentin et al., 1996; Haxby et al., 1996; Liu et al., 2002, 2000). The grand average ERF over all subjects is shown for a voxel situated in the right FFA in Figure 6.5. Both the TASER and statically derived traces are shown. Reassuringly, the plots are very similar, although subtly different. This instils some confidence that the temporally adaptive philosophy is not detrimental to recovering any underlying brain activity in the time dimension. The sensor level and HMM-beamformer ERFs are plotted for a comparison in Figure 6.6.

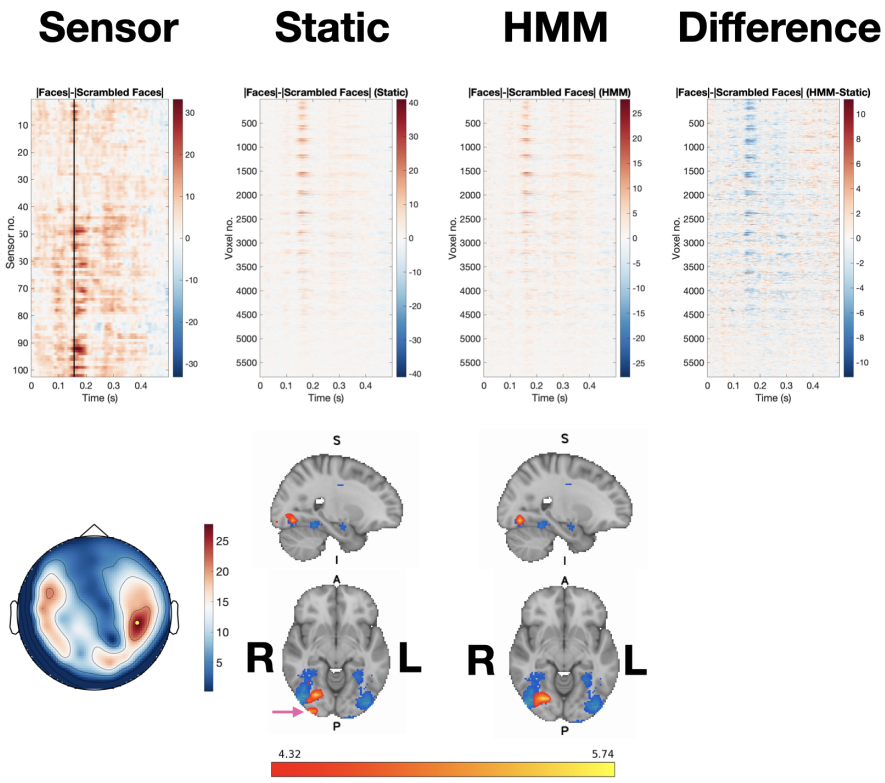


Figure 6.3: Contrasting stationary and non-stationary beamformer results. Top: Evoked response analysis averaged over all subjects and all sessions. From left to right, the averaged (absolute) evoked response is shown for sensor level data, a static beamformer, a HMM beamformer and the difference in the source space images is shown. As expected, a clearer and cleaner M170 response is seen in the source space data. Non-stationary beamformers give a different result to the conventional static approach, with most voxels showing a decreased difference between the two conditions (in terms of absolute magnitude). Bottom, left: topography at the +168ms latency for the averaged Faces-Scrambled Faces at the sensor level. The largest sensor change is marked in yellow. Bottom, right: Uncorrected t-statistical maps at +168ms for the static beamformer (left) and HMM beamformer (right). The red-yellow map shows the statistical values; the blue-green colourmap shows the Z-scored meta-analysis of Fusiform Face Area studies (downloaded from Neurosynth). The HMM map peaks over the right fusiform face area, as expected. The static beamformer image peaked outside of the brain and is more dispersed, involving early visual areas, too. Both maps are thresholded at the 75th percentile.

Latent Variables Correlate with Task Structure

Figure 6.8 shows how both unsupervised inference techniques correlated with task structure and the brain's response to stimulus presentation. It should be re-iterated that neither the HMM nor VRAD are ever directly exposed to the timings of the stimulus onset. The top row shows the evoked state-occupancy after post-hoc epoching for the HMM. Two prominent states, 3 and 13, are selectively presented. The former appears to be an early (M100) visual response to the initial presentation of the face, whilst the latter peaks later at around 200ms over more lateral sensors.

Rows 2 and 3 show the same evoked waveforms, now of the (trace-weighted[1]) mixing coefficients, α_t, for a $J = 10$ and $J = 20$ model. Note that only the higher

[1]The trace of the corresponding spatial map is used to multiply the mixing coefficients so that the presented timeseries is somewhat representative of the scale of the inferred variance.

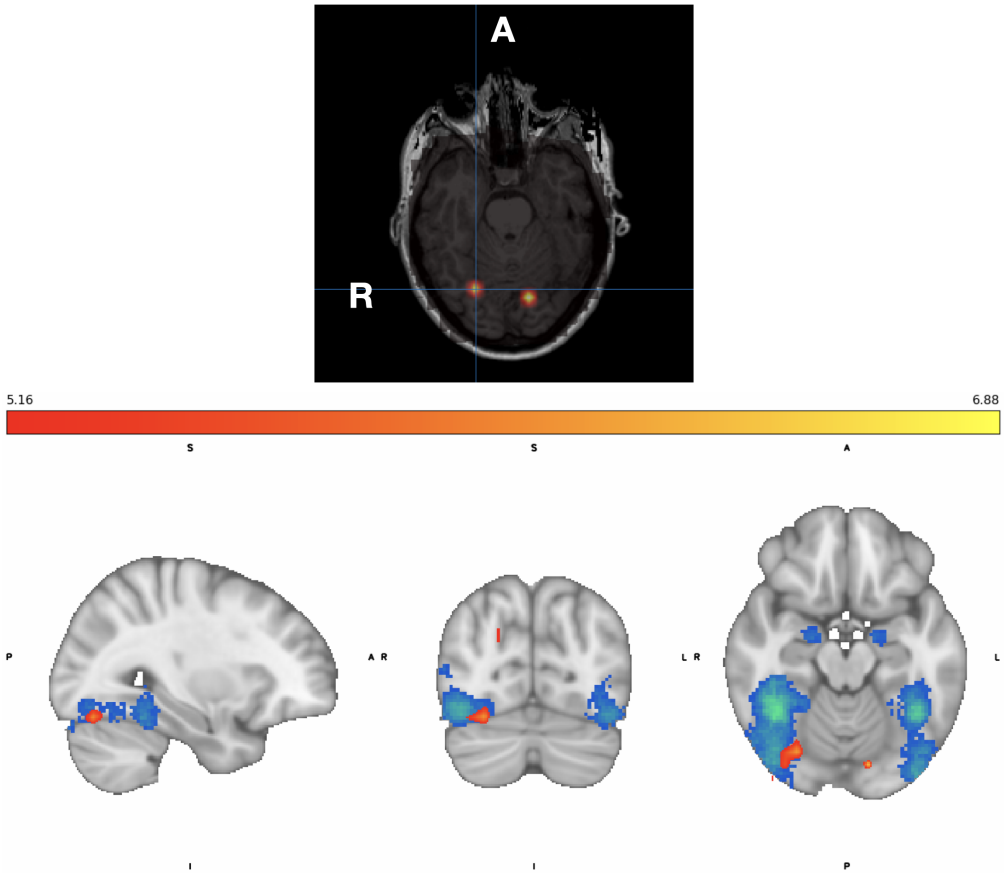


Figure 6.4: VRAD T-stats at 168ms (famous - scrambled faces). Top: positive clusters which survived the multiple comparisons correction in subject space. Bottom: statistical maps (red/yellow) for the contrast thresholded at the 75th percentile overlaid onto the Neurosynth meta-analysis result (blue/green). As anticipated, the map peaks in the right FFA, but also in the left lingual gyrus.

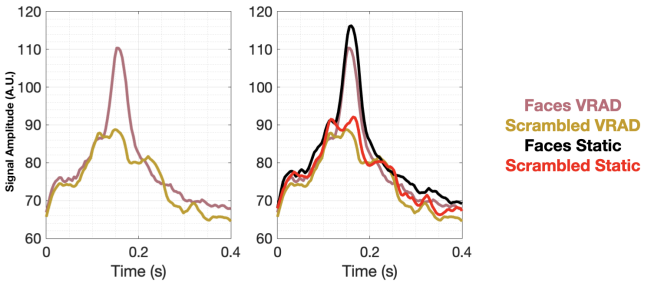


Figure 6.5: TABFER recovers the classic M170 response in the right FFA. Left: mean of the (absolute) evoked response to the presentation of faces (purple) and scrambled faces (gold) across all subjects when using the TABFER. A large deflection is seen at 170ms post stimulus onset, as expected. The same result is compared with the static method on the right. Traces are very similar, with the static method obtaining a higher amplitude in the faces condition, but with a larger response in the scrambled trials.

order model was used in the source reconstruction. Like the HMM result, there is a notable shift in the traces at around 100ms (especially in the order 10 model). However, as the VRAD model is non-categorical, these evoked traces are not as prominent, and it is the sum of all of the individual components which is important.

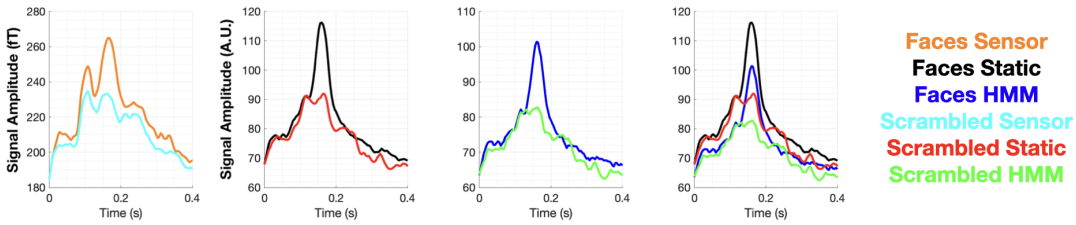


Figure 6.6: Evoked response traces at the sensor level, and at the peak statistic voxel in the FFA (taken from the HMM map) in the source space. Note the subtle changes in the ERF traces and the decreased magnitude in the HMM result. Source traces are plotted on their own scale and together to aid visual comparison.

This is shown in the right hand columns at 168ms for both model orders (the mean $\mathbf{C}_t$ matrix over all trials is plotted for this timepoint), where power is found over left lateral sensors.

Figure 6.7 shows how both methods are capable of learning covariance matrices which resemble artefactual (e.g. $J = 20$, component 2 - eyeblink) and neural (e.g. $J = 20$, component 10 - motor response, likely to the button press) topographies.

Discussion

We have shown in this chapter that TASER can be applied to more challenging datasets and provides meaningful results. Both dynamic beamformers provided cleaner and clearer foci in the right FFA (Figures 6.4 and 6.3). We reconstructed and recovered the classic M170 (Figure 6.5) and showed that, in this dataset, using a large number of weights did not drastically distort signals of interest in the time domain.

Here we ran VRAD and the HMM separately for each of the 6 sessions. This is in contrast to the approach taken by Quinn et al. (2018). We decided not to concatenate the trials over sessions to avoid the risk learning components/states which changed purely due to different head positions (although the MaxFilter operation should go a long way to reduce this affect).

Face perception is known to evoke and induce responses at different latencies and frequencies (e.g. the M100 and M220) other than the one analysed in this Chapter (Perry and Singh, 2014; Quinn et al., 2018; Wakeman and Henson, 2015). Quinn et al. (2018) reported a statistically significant change in state occupation for faces - scrambled faces at latencies between 500 and 800ms in the frontal lobe. Our simulations suggest that there may have been a more pronounced difference between the two reconstruction methods (static versus dynamic) at these brain regions, due to the relatively weak lead fields attained in a cryogenic MEG system. We purposefully limited our analysis in this work as we expected to see the biggest signal (change) in the faces minus scrambled faces contrast.

In contrast to Chapter 5, here we took the grand average of all source space data across all sessions and subjects. There is a strong possibility that any subtle changes to the source space results may have been dampened or outright removed from the act of averaging over this many datasets and trials, with the heterogeneity over subjects removing any differences afforded by the dynamic technique.

The preprocessing carried out prior to inference in this chapter was coarse. We did not follow recommended best practice (Gross et al., 2013), nor did we execute any ICA as was done in Quinn et al. (2018). Specifically, we did not carry out

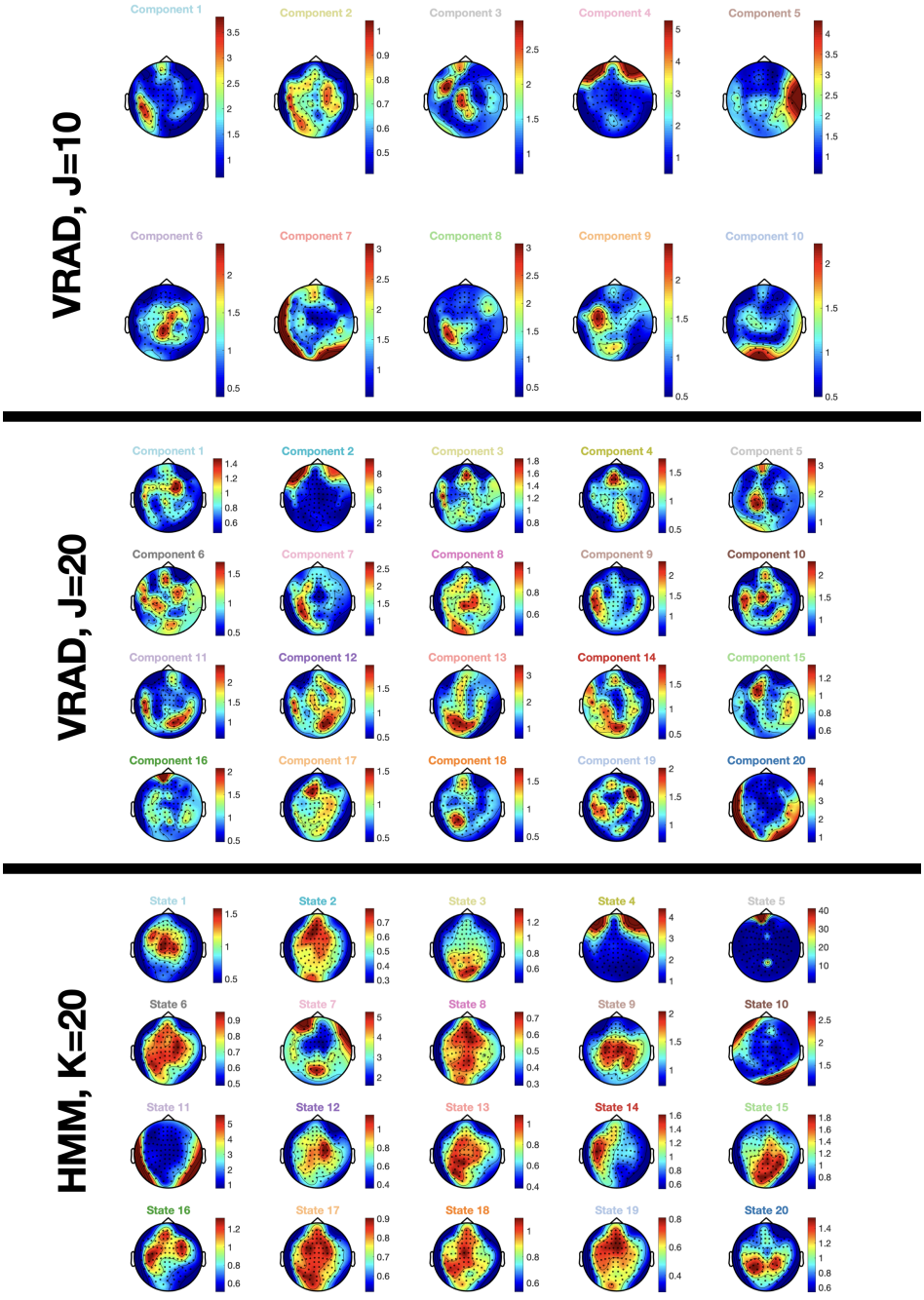


Figure 6.7: Various components/states learnt by VRAD and the HMM. 10 and 20 component models were run for VRAD to assess how mode number effected reliability. A 20 state HMM was run as a comparison. Also see Figure 6.8.

any rigorous data cleaning in the form of channel or artefact/trial rejection. This includes not executing any ICA to remove heartbeat/ocular components from the sensor level data. These components were however observed in both the HMM and VRAD states/components (Figure 6.7), potentially at the expense of more interesting neural phenomena. Future work could look at how components inferred by VRAD could be used in preprocessing operations.

In spite of these shortcomings, our source localisation was successful, even though we used the (typically) more noisy magnetometer channels. Beamformers are famous for their ability to suppress activity originating from outside of the forward

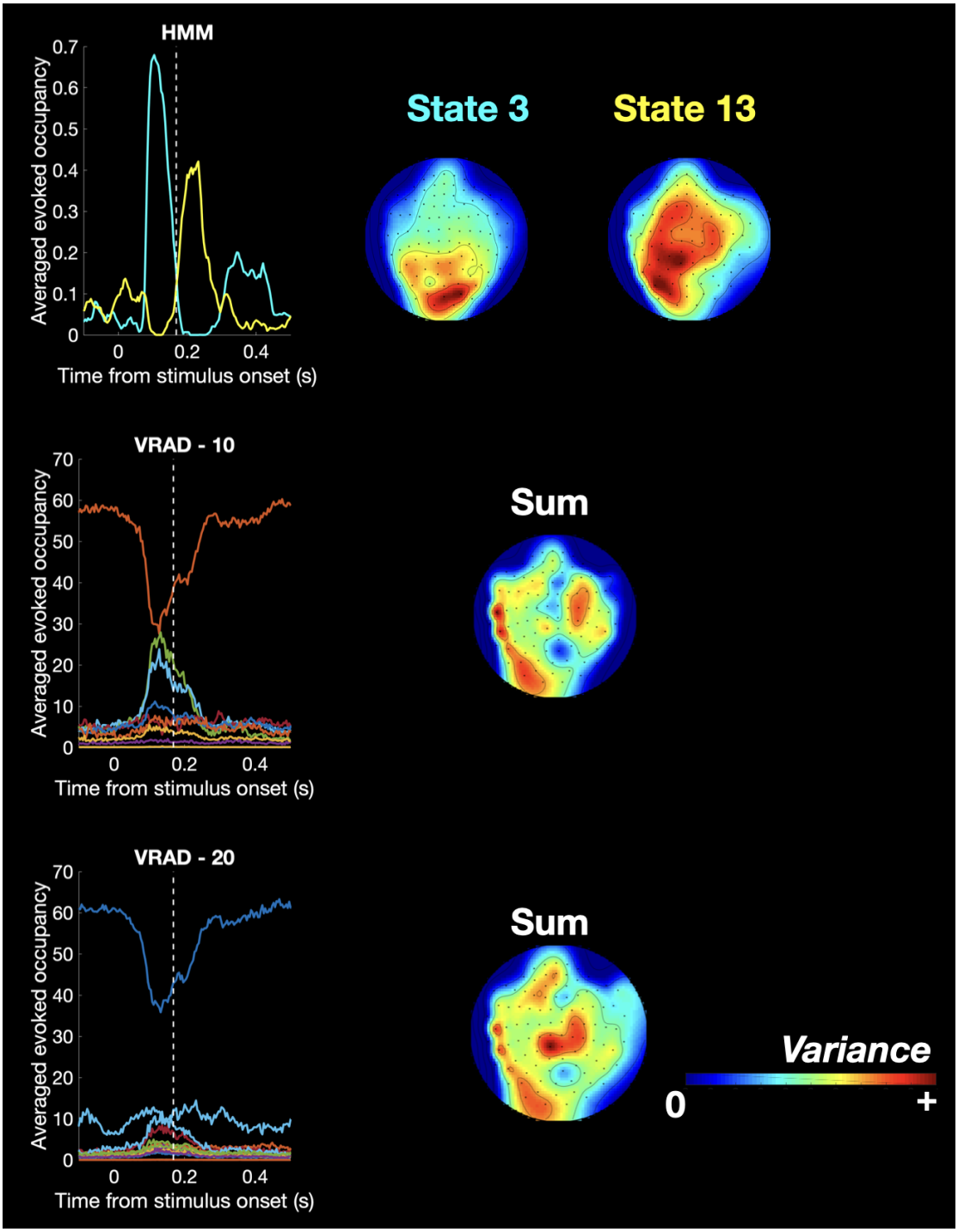


Figure 6.8: Visualising the learnt spatio-temporal sum by VRAD and the HMM. Top: the two most prominent evoked HMM states are shown. An early visual response is seen at around 100ms (state 3), followed shortly after by a more anterior lateral response onsetting at about 170ms. The sum derived from the 10 and 20 element VRAD models are shown for 168ms (marked with the dashed white line on the traces). Note the similar result between the two model orders. A clear evoked response is present in the $J = 10$ model at 100ms, again probably corresponding to the M100 response, with the orange component becoming recruited less at this time. The VRAD mixing coefficients were weighted by the trace of their associated spatial map. No baseline correction was applied to these traces.

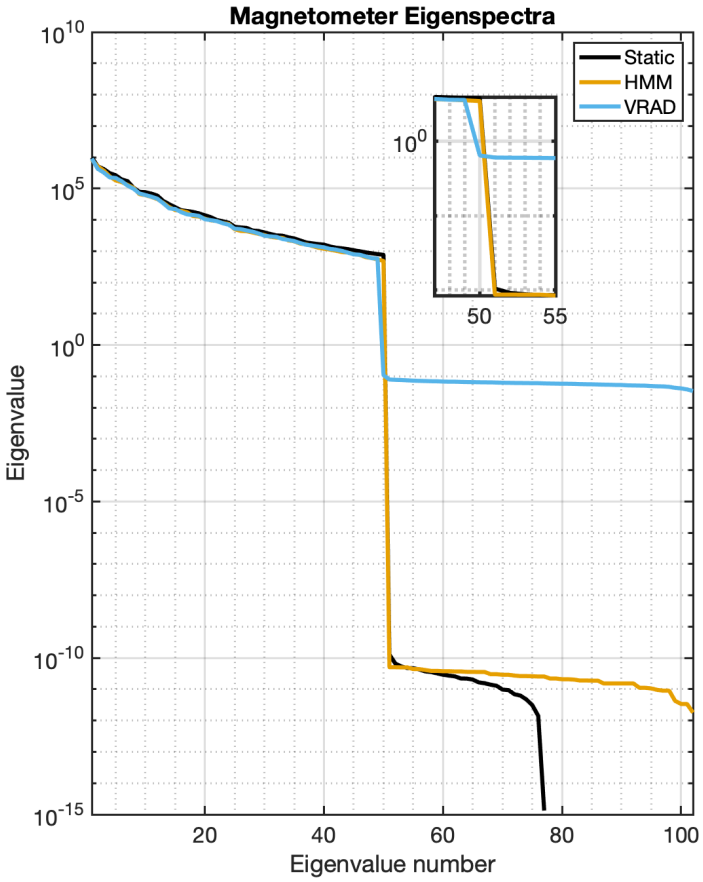


Figure 6.9: Eigenspectra for the differently derived covariances. Static, HMM and VRAD inferred matrices are shown in black, gold and blue, respectively. Note the large drop-off at around 50 components as a result of the MaxFilter operation. This is highlighted in the enlarged inset plot. The VRAD eigenvalues do not drop to zero due to the numerical correction added during inference.

model (see Litvak et al. (2010) and Boto et al. (2021) for examples). Future projects could examine the extent to which using time-point specific covariance matrices improve/reduce the efficacy of this suppression over static methods. Now that we have demonstrated TASER's ability to source localise deep sources, it would be interesting to apply the method to subcortical structures such as the hippocampus, where using finely tuned reconstruction weights could stand to significantly improve the reconstruction accuracy.

The data studied here were susceptible to coregistration error. Use of the MaxFilter software to correct for head motion and changes in position across sessions will have gone some way to mitigate this affect, yet ultimately each subject's forward model will have been imperfect. In this spite of this, both static and dynamic beamformer methods did successfully localise the FFA, with the dynamic beamformers giving slightly more accurate results.

An outstanding question revolves around this coregistration error and its impact on attempts to use more than one set of reconstruction weights. Future work should quantify how error in the lead field matrix acts to diminish improvements afforded by using TASER.

Finally, Figure 6.8 serves as a stark reminder that the added modelling capability gifted to us by VRAD can come at the expense of immediate interpretability. In this thesis we use the elements learnt by VRAD as a means to an end - to define the dynamic covariance matrices.

Conclusion

This chapter has shown that using more than one reconstruction weight for MEG source reconstruction can be beneficial even in the presence of coregistration error, *and* when one is trying to reconstruct activity in deep locations. Our results show that the approach can translate, and scale, to different scanner types and larger datasets, respectively.



7 Discussion and Conclusion

Lessons from this thesis

The primary goal of this thesis was to investigate whether pre-existing approaches to source imaging can be improved (or hindered) by incorporating dynamics into the modelling of sensor level data. The secondary goal of this degree was to develop a framework to enable this temporal modelling, with increased flexibility over established techniques (Baker et al., 2014).

The work developed in this thesis built upon an appreciation for the approaches to source reconstruction as laid out in Chapter 2. The common limitation amongst these algorithms was identified: data covariance is often implicitly assumed to be static.

To ameliorate these historically successful algorithms, we first executed a literature review into options for dynamic covariance modelling. This was presented in the first half of Chapter 3. Inspired by a principled and powerful generative model - the Hidden Markov Model Rabiner and Juang (1986) - we sought to develop a new tool for multivariate neuroimaging timeseries data, designed with the specific application of source localisation in mind. This meant that it had to be capable of modelling dynamic covariance as an ensemble of components, rather than a categorical model, if a temporally adaptive implementation of multiple sparse priors was to be realised (Friston et al., 2008).

This required the uptake of stochastic inference techniques to circumvent the need to derive update equations which may have proven to be intractable. This inference scheme also allowed us to model timeseries data with more sophisticated observation models. This facilitated the use of a highly non-linear generative model in the form of a LSTM (Hochreiter and Schmidhuber, 1997). We showed in simulation (Figure 3.7) that this novel method captured longer range temporal dependencies in the data, something to which the HMM was blind.

By taking lessons from the ML community (Kingma et al., 2015; Kingma and Ba, 2014), we were able to develop and deploy an efficient model where the number of trainable parameters was independent of the length of the dataset, making it immediately scalable to much larger datasets. This *amortised inference* used bi-directional LSTMs to parameterise a mapping from MEG sensor data to the latent variables (the mixing coefficients, α_t). Importantly, whilst we treated the ML architecture and its components as a bit of a black-box, we started by writing down a generative

model and used the Variational Free Energy as a cost function (Friston et al., 2007). This allowed us to learn parameters of interest within a well-defined framework, with a penalty term in the loss controlling for complexity. This unsupervised learning algorithm was shown to be more accurate at reconstructing dynamic covariance than the more restrictive HMM when the ground truth generative process was not categorical (Figure 3.8).

The concept of Temporally Adaptive SourcE Reconstruction for MEG was presented in Chapter 4. Past dynamic localisation algorithms were discussed before the novel, completely time-varying method was described. This philosophy was motivated by a realistic set of simulations (Figures 4.3-4.9) which showed significant improvements in temporal reconstruction accuracy at deep and radial brain regions when a dynamic reconstruction method was adopted over a static one. Our literature review of previous non-stationary attempts told us that improvements also stood to be gained when attempting to reconstruct activity at neighbouring sources in the brain (Woolrich et al., 2013; Dalal et al., 2008; Barratt et al., 2018).

TASER was deployed on a high precision MEG dataset (Barratt et al., 2018) in Chapter 5. We demonstrated the use of the TAMSP and TABFER versions of the technique and made comparisons with pre-existing static counterparts. We began this chapter by showing how different assumptions revolving around temporal stationarity reveal different properties of the same data (Figure 5.3), with both unsupervised methods (the HMM and VRAD) revealing a highly dynamic picture of the spatial variance.

Using a generative model source localisation technique (TAMSP), we were able to successfully delineate the origin of spontaneous neural rhythms arising in primary motor cortex and found peak statistics at the anatomically defined “ground truth" locations. We also showed that the latent dynamics closely married with findings from the beta burst literature, a new hot topic in the electrophysiological community Seedat et al. (2020); Quinn et al. (2019). These findings were repeated for the spatial filter implementation of TASER: Figures 5.5 and 5.6 show that an increasing amount of temporal components (manual refinement of reconstruction weights, $K = 11$ and $K = 20$ HMM models all the way up to 400 different reconstruction weights derived via VRAD) were needed to arrive at the classic homunculus. As expected, images became more resolved as the accuracy of results improved (in the sense that there was less spatial overlap after thresholding). We also saw a statistically significant shift in the peak statistic location for the dynamic beamformers, with results being in closer alignment with fMRI findings obtained from a similar paradigm (O’Neill et al., 2020).

Importantly, these improvements in source localisation stemmed from the spatial and temporal elements identified at the sensor level by the unsupervised inference regimes. These features resembled bodily function (heart and ocular function), as well as neural phenomena, as demonstrated in Figures 5.7 and 5.8. Figure 5.9 highlighted the advantages of taking a time-varying perspective of the data, where beta bursts were revealed and were shown to exhibit substantial inter-trial variability. Similar spatial and temporal properties were found with the HMM, a tried and tested technique, boosting the confidence in our results. Further validation was gained when we used one of the advantages of the HMM - that it is a categorical state space model - to derive state specific source power maps, alongside temporal statistics. These are shown in Figure 5.12. The former analysis confirmed our hypotheses that the sensor level maps for states 4 and 14 pertain to periods of beta power suppression, whilst states 15 and 19 were beta rebound events. The latter analysis gave us reason to have faith in our covariance estimation. A similar attempt to quantify the number of timepoints used in the calculation of the VRAD maps was attempted.

Chapter 6 saw us turn to a more conventional multi-subject multimodal dataset (Wakeman and Henson, 2015). We saw that TASER was able to successfully identify the FFA in the faces - scrambled faces contrast. Whilst the static result peaked outside of the head, both non-stationary beamformers (HMM and VRAD) peaked in the right FFA. Our results showed that TASER faithfully recovered the M170 response, and that this was very similar to the statically derived traces. On the one hand, this result can be interpreted as underwhelming. On the other, a more optimistic conclusion is that this is a validation of the method - it is clear that our use of more than one reconstruction matrix was justified in this context, and that this was not detrimental to the reconstruction of the timeseries, and the resulting well-documented evoked response. We should however stress that we carried out a very limited analysis of this dataset and that improvements/degradations may have been observed at different latencies and/or in induced responses (Perry and Singh, 2014).

We were reminded in this chapter that VRAD is a non-categorical model. This can sometimes make the interpretation of individual elements hard. This is shown in Figure 6.8 - whilst there is a clear response at 100ms in the VRAD approach, this is not as clear as the HMM result, where a clear early visual response can be seen in state 3. Nevertheless, when summed together, there is a clear structure in the sensor level variance in the VRAD models. Spatially, the VRAD elements again encoded physiological processes (heartbeat, eyeblinks and motor responses).

Taken together, we have shown in this thesis that appropriate unsupervised learning techniques can facilitate TASER. This in turn can provide more accurate and resolved images of brain activity in the spatial domain, whilst maintaining key temporal properties of the data. The peripheral results learnt along the way often correlate with neural processes and the task structure, and may be of interest to the enthusiastic neuroscientist.

Now that we have outlined the main learnings from this doctorate, we turn to some discussion points common to all chapters, and take a look to the future of the M/EEG inverse problem and applications of TASER.

The struggle with TAMSP in its current implementation

In our implementation of TAMSP, we packed several hundred priors into the VRAD model, i.e. $K \approx 300$ in Equation (4.2). This value of K was limited by the memory afforded by the TensorFlow architecture underpinning VRAD: we increased this value until we reached the size of the RAM on the GPUs.

This made obtaining high spatial detail for whole-brain TAMSPs impossible in a general context. Instead, as shown in the work presented here, we opted to tightly pack our patches into predefined ROIs in the case of the motor experiment, or used a whole-brain parcellation to obtain whole-brain reconstruction for the Wakeman and Henson dataset, but at the expense of spatial detail (these results were not presented in this thesis). Still, this did allow us to use K different covariance patches, as opposed to just 1 as in the classic MNE approach. We found that VRAD reduced the size of the model to the rank of the data (again, not shown here), a point we shall return to in just a moment.

On reflection, trying to model the covariance dynamics at the same time as doing the source reconstruction was needlessly computationally inefficient. To circumvent the issue of getting detailed whole brain activity, whilst accounting for the dynamics and gaining the advantages of working in a generative model based reconstruction pipeline, a three-step process is a likely way forward:

1. A handful of components are inferred at the sensor level, without restriction, as in the TABFER flavour of the algorithm.

2. K-means clustering (or some other method) is used to reduce the temporal dimension to a more manageable number of snapshots of covariance.

3. A static implementation of MSPs is fitted to the local approximation of $\mathbf{C}_Y^t$. This is repeated for each of the matrices derived in step 2.

Crucially, by switching TAMSPs to this regime, we will be able to tune the spatial resolution of the reconstruction weights depending on the instantaneous properties of the data. This is achieved by using a Fisher-scoring algorithm to reduce the number of patches from several hundred to just a handful, as redundant components are either combined or removed (Friston et al., 2002). This is a sensible (and indeed necessary) step implemented in the static version of MSP (Friston et al., 2008). By carrying out this patch reduction on the local estimates of the dynamic covariance, we should be able to obtain the best of both approaches - accounting for radical evolutions of the covariance in time, whilst not fitting an obnoxiously large number of redundant priors to the covariance over the duration of an entire experiment.

If this approach is successful, it will facilitate higher spatial detail even when there is no strong hypothesis, and opens up TAMSPs to be used on resting state scans of arbitrary length, a major advance on the current temporally static algorithm.

Finally, whilst the rank of MEG data is limited by the number of sensors in our array and certain preprocessing steps, the proposed hybrid approach means that the degrees of freedom in the data can be adaptively spread around the brain. This is in contrast to the current implementation of TAMSPs, where once a brain region is “knocked-out" by the inference, it cannot be adaptively added back into the model.

Some arguments against TASER

Having presented some of the potentials of the methods developed in this thesis, we must acknowledge that there are some arguments which can be made against the use of multiple inversion weights. We shall explore these briefly now.

Claim: using multiple weights obfuscates source space images

It can be shown with relative ease that the variance of a beamformer reconstructed dipole at location q' is given by

$$\hat{\sigma}^2_{q'} \propto \sigma^2_{q'} + \mathbf{w}_{q'}\left(\sum_{q\neq q'} \mathbf{h}_q \sigma^2_q \mathbf{h}_q^T\right)\mathbf{w}_{q'}^T + \sigma^2_\epsilon \mathbf{w}_{q'}\mathbf{w}_{q'}^T, \tag{7.1}$$

where $\hat{\sigma}^2_{q'}$ is the approximated variance, $\sigma^2_{q'}$ is the underlying true variance and $\mathbf{w}_i$ and $\mathbf{h}_i$ are the beamformer weights and lead field vector at location i, respectively. In words, the latter two terms in Equation 7.1 represent the variance contribution due to source leakage and (white) sensor level noise. Focussing on the source leakage term, one can therefore see that the reconstruction error is a function of the weights for dipole q' and the lead fields from the rest of the brain.

Now consider the (noiseless) case where we have just two different sets of beamformer weights for conditions A and B. The error due to signal leakage for these two conditions are

$$\Delta A = \mathbf{w}_{q',A}\left(\sum_{q\neq q'} \mathbf{h}_q \sigma^2_{q,A} \mathbf{h}_q^T\right)\mathbf{w}_{q',A}^T, \tag{7.2}$$

and

$$\Delta B = \mathbf{w}_{q',B}\left(\sum_{q\neq q'} \mathbf{h}_q \sigma^2_{q,B} \mathbf{h}_q^T\right)\mathbf{w}_{q',B}^T, \tag{7.3}$$

where $\sigma_{q',A}$ and $\sigma_{q',B}$ are the ground truth variance at vertex q' during condition A and B, respectively. The argument can be made at this stage that if we were to create some image between the two conditions

$$\hat{\sigma}^2_{q',A} - \hat{\sigma}^2_{q',B} = \sigma^2_{q',A} - \sigma^2_{q',B} + (\Delta A - \Delta B), \tag{7.4}$$

then we would not know as to whether our image was a result of the true effects, or due to different spatial leakage profiles arising from the use of different beamformer weights.

The alternative is to use one average set of reconstruction weights, calculated from activity across conditions, $\mathbf{w}_{q'} = \mathbf{w}_{q',AB}$. The hope here is that the error terms due to source leakage will cancel out. Taking this idea forward, the image is given by

$$\hat{\sigma}^2_{q',A} - \hat{\sigma}^2_{q',B} = \sigma^2_{q',A} - \sigma^2_{q',B} + \mathbf{w}_{q'}\left(\sum_{q\neq q'} \mathbf{h}_q \sigma^2_{q',A} \mathbf{h}_q^T\right)\mathbf{w}_{q'}^T - \mathbf{w}_{q'}\left(\sum_{q\neq q'} \mathbf{h}_q \sigma^2_{q',B} \mathbf{h}_q^T\right)\mathbf{w}_{q'}^T \tag{7.5}$$

$$= \sigma^2_{q',A} - \sigma^2_{q',B} + \mathbf{w}_{q'}\left(\sum_{q\neq q'} \mathbf{h}_q \sigma^2_{q,A} \mathbf{h}_q^T - \mathbf{h}_q \sigma^2_{q,B} \mathbf{h}_q^T\right)\mathbf{w}_{q'}^T, \tag{7.6}$$

$$= \sigma^2_{q',A} - \sigma^2_{q',B} + \mathbf{w}_{q'}\left(\sum_{q\neq q'} \mathbf{h}_q \mathbf{h}_q^T \left[\sigma^2_{q,A} - \sigma^2_{q,B}\right]\right)\mathbf{w}_{q'}^T. \tag{7.7}$$

We can therefore see that the change in the image due to leakage does not cancel out, even if we use a mutual set of reconstruction weights, so long as the summation above does not equal zero. This will generally be the case since the variances tend to change between the two conditions, $\left[\sigma^2_{q,A} \neq \sigma^2_{q,B}\right]$. It is therefore unlikely that this argument can be correctly made against TASER.

On distorting reconstructed signals in time

Another fear stemming from the use of multiple reconstruction weights revolves around the underlying waveforms of neural activity. Consider a hypothetical situation where a ground truth signal has a sinusoidal profile. In the static reconstruction case, the estimated source activity is given by a *single* weighted linear sum of the sensor timeseries. There is a risk that by introducing dynamics into this weighted sum that we could distort the underlying signal. Figure 7.1 depicts this concept.

Whilst a very reasonable concern, we saw both in simulation (Chapter 4) and in real data (Chapter 6) that our dynamic implementations did not distort the underlying neural timeseries. In contrast, we were able to achieve better reconstruction accuracies, and recovered the classic and well-documented M170 evoked response. This may be down to the fact that, whilst our approach was dynamic, we used a relatively small number of reconstruction weights (400) relative to the number of timepoints in the experiment.

Concern: what if we derive erroneous covariance matrices?

A final concern with the technique[1] centres around being “wrong" in our estimates for the covariance. Put another way, by trying to be “too clever", using more than one set of reconstruction weights, we risk tripping over our own ambition. This argument certainly has merit when data have a particular set of characteristics. Attempts have been made to provide a thorough description of these properties throughout this work - the signal bandwidth, SNR, the number of data points available and the physical relationship between sensors and sources (encoded in the form of

[1] Which we shall entertain in this document - there are, of course, likely to be others!

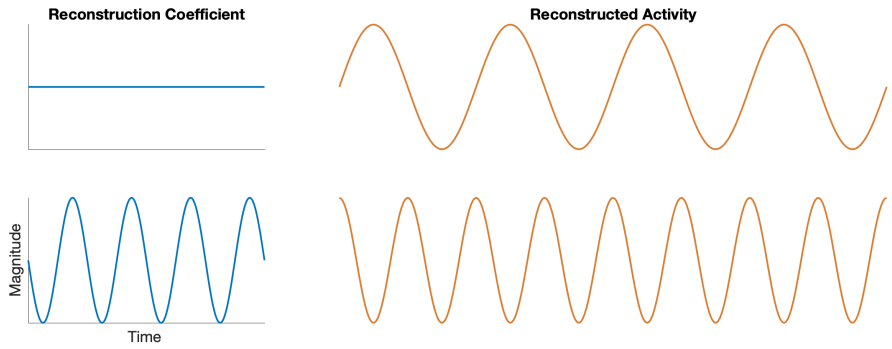


Figure 7.1: One worry about using multiple conversion weights (as opposed to just one) is that the dynamics in the weights themselves could introduce error into the reconstructed timeseries. This idea is depicted here. Top row: a set of static reconstruction coefficients/weights will not alter the dynamics of the sensor level signal (only their relative weighting across all time). Bottom row: in contrast, if the weights were to wrongly begin oscillating, then they could introduce error into the ground truth signal. We addressed this concern in both simulation and real data in this thesis.

the lead field matrix) are all implicated. Perhaps the most simple, yet important, take home message from this thesis is that in order to realise the advantages of TASER (or the dynamic modelling of MEG data, more generally), we need adequate amounts of data, as machine learning methods (the HMM, VRAD, or other) are generally “data hungry" approaches.

To address some of these concerns, we showed in simulation that by introducing a non-categorical model (VRAD) that we were better able to reconstruct dynamic covariance profiles (Figure 3.8) over a discrete state-space alternative (the HMM), thus reducing the covariance reconstruction accuracy. We were also able to show empirically that we obtained at least parity, if not an improvement in source reconstruction accuracy across spatial (c.f. Figure 5.5 and Figure 6.4) and temporal dimensions (Figures 4.4 and 6.5). Away from source localisation, our inferred latent variables closely aligned with task structure (despite the fact that we used unsupervised techniques) and classic spatial topographies well documented in the M/EEG literature. Attempts were made to quantify the confidence in our estimates for covariance modes/states throughout this work. Future work can look at and quantify how covariance model error translates into source reconstruction accuracy when a dynamic model is used.

What VRAD gives us

VRAD provides a detailed breakdown of latent processes occurring across space (Figures 3.8, 5.7 and 6.3) and time (Figures 3.8, 5.8 and 6.6) originating from both inside and outside of the brain, or equivalently, a depiction of signals of interest and sources of noise. Here we have shown that these spatio-temporal elements show task relevance and that this description of the data ultimately leads to more physiologically agreeable source reconstruction results over the static beamformer counterpart.

The complementary perspectives of the dynamics of covariance portrayed in Figure 5.8 provide different insights into the data and could be useful for different applications. In this study, the artefactual components give information as to when MEG data is confounded in the continuous time domain, and in future could be used to segregate, regress or remove bad segments of recordings. Conversely, the epoched data

helps identify task-specific dynamics. This classification of components is neither a primary (or even secondary) goal of the TASER framework but comes “for free" with VRAD.

Hyperparameter Choices

A number of hyperparameter choices were made in this work. We discuss these briefly here.

Choosing J: The number of modes

VRAD has many hyperparameters which must be specified *a priori* (see Chapter 3). Of these parameters, the perennial choice of the number of modes[2], J, must be addressed. Speaking heuristically, the higher the value of J, the more granular a picture we can anticipate in the covariance dynamics, and hence the greater variability in the source reconstruction weights, provided that we have a large enough number of data points, and that are indeed dynamics to be recovered. However, as stressed earlier, when tackling the inverse problem with a beamformer, there is a trade-off between the number of modes used and the accuracy of source inversion weights which must be considered (Brookes et al., 2008).

One can adopt formal strategies for choosing the optimal number of modes, such as finding the model which maximises the Free Energy and/or turning to cross-validation methods. We find however (both in soon to be published work and in a very similar model (Skates, 2019)) that the Free Energy increases monotonically with the value of J. An alternative approach is to start with a small number of components and increase these until almost identical results for the time-varying weighted sum (i.e. the left hand side of Equation (3.4)) are obtained.

It is worth reiterating at this stage that for source localisation applications, we are not necessarily concerned with interpreting the $\mathbf{B}_j$ spatial modes or the dynamics, meaning that we are able to accept some duplicity in the inferred latent variables. Moreover, the reader is reminded that it is ultimately the final linearly weighted sum of terms which determines the instantaneous covariance estimation, as opposed to any of the components considered individually.

Here we chose $J = 20$ components for the VRAD model, as we found that this represented a good compromise between robustness and flexibility in the spatio-temporal dynamics of the covariance for this particular dataset. Another reason for choosing such a (relatively) high number of modes is that previous dynamic source localisation results reported little to no difference in the inversion results when a small number of components were used (Woolrich et al., 2013). Finally, by matching the number of modes to the number of states used in the HMM model we are able to draw qualitative comparisons between the two techniques.

Choosing K: The number of centroids

Another important hyperparameter for TASER is the number of clusters that we use to discretise the alpha values. If we choose too small a number then we will misrepresent the true dynamics and are at risk of introducing error into our dynamic source reconstruction; choosing too high a number significantly increases the computational burden placed upon us. What K puts us in the goldilocks zone?

To assist with choosing a practical number of clusters we plotted the percentage error between the discretised and the original alphas for the motor headcast dataset (bottom of Figure 5.13). We (somewhat arbitrarily) introduced the criterion of achieving

[2] Or states, K, if one is using a HMM.

less than a 10% mean percentage error, and settled with using $K = 400$ centroids. Whilst the feat of calculating 400 sets of weights is still quite computationally demanding, we believe that it represents a pragmatic balance between losing accuracy in the estimation of dynamics (or just using a static model altogether) and trying to calculate a set of reconstruction weights for every point in time. The optimal choice for K could be explored more in future work via simulation.

The above two paragraphs are another way of asking “just how dynamic can we really be?". We took changes in $\mathbf{C}_Y(t)$ at face value in this thesis and used each different instance of this matrix to give a set of reconstruction weights. In the future a more statistically driven approach could be adopted, i.e. only covariance matrices which are some criterion away from the time-averaged covariance are used to give different weights, thus reducing the risk of using too many different matrices.

VRAD vs the HMM

The convenience of a state-space model

The abstraction of representing data as a set of mutually exclusively activating states in the HMM turns out to be very convenient for post-hoc analyses and the calculation of state-specific power maps. This is evidenced in the state time course in Figure 5.11B, where at nearly all points in time one state dominates the time course. It is therefore a safe approximation to use the Viterbi path (i.e. the argmax of the inferred probabilistic state timecourse) to represent the data dynamics. This facilitates the pooling of data points from non-contiguous segments into different windows for the calculation of statistics with there being little to no ambiguity as to which state a particular data point is attributed.

The increased flexibility in temporal modelling introduced by VRAD means that post-hoc analyses are not as straight forward to carry out. The continuous α_t traces shown in Figures 5.8 and 6.8 show how a mixture of components simultaneously contribute to the instantaneous covariance. There are two distinct challenges that this mixture poses:

1. **The Calculation of Mode Statistics**
   As the mode activations are continuous, there is no sense of them switching “on" and “off" as in a HMM state time course. If one is to derive modal temporal statistics, we must hence make an effort to classify when a mode is active.

   One option is to fit a Gaussian Mixture Model (GMM) to the mixing coefficients α_t and learn one distribution attributed with the mode being “on" and another as to when it is “off". This binarisation of the mixing coefficients gives something akin to a Viterbi path, which allows one to calculate statistics analogous to those plotted in Figure 5.12.

   Figure 5.14 shows an implementation of this, with the top panel showing the result of fitting a 2 component GMM to the mixing coefficients. Red curves show data points where the component was approximated as being “off", blue lines indicate when “off". These data were first multiplied by their corresponding trace (as in Figure 5.10), re-normalised (so that the resulting trace-weighted alphas summed to 1 at each time point) and then transformed to a more gaussian distribution by using a reciprocal transform. The bottom panel shows the resulting amount of time that we consider a mode to be “on" (left), and an example 2 second segment of the (transformed) $\alpha_{1,t}$ trace alongside the derived threshold value.

   Whilst elegant, this only gives a crude approximation to how much a component is “on" over the duration of an experiment. Nevertheless, it is useful in quantifying

how much confidence we can place in our covariance estimation. The quality of this estimate is ultimately dependent on the GMM fitting performance. It is worth noting that the distributions vary considerably depending on the nature of the component: artefactual components (e.g. 4, heart beat) are very bimodal whilst neural elements (e.g. 10 and 11) have tighter distributions. This variation makes it hard to choose the optimal number of components in the GMM and generalise the preprocessing steps (e.g. outlier removal) for the mixing coefficients. For simplicity, we opted to use a 2 component model. This could be made more systematic (improved) by using some form of cross-validation to determine the optimal component number. Another way of increasing the rigour of this estimation would be to incorporate the dynamics of the mixing coefficients themselves into this classification. A natural extension, ironically, would therefore be to apply a 2-state HMM to these data. Finally, the statistics shown in the bottom left of Figure 5.14 are calculated after hard classifying the components into on and off. A more representative set of values may be derived by using the probabilistic weighting from the GMM result (c.f. Zhang et al. (2000)).

To try and define the uncertainty of these timings we also turned to a more basic approach, using Otsu's method to binarise the timecourse. These two methods combined gave us an upper/lower bound for the component activations.

2. **The Construction of VRAD power maps**
   The HMM state time course gives a set of timepoints that can be used to easily calculate state power maps (Figure 5.12). Maps could be constructed in a similar manner for VRAD by fitting a linear model to the source space spectral power (e.g. wavelet transformed data), using the mixing coefficients as regressors.

Despite these practical hurdles (which certainly require more attention in future work), we saw in this thesis that VRAD does give a more dynamical picture of the variance. An example of this is shown in Figure 7.2. Here, the instantaneous covariance as modelled in VRAD was reconstructed for 25 random points in time during which a HMM state was activated. Overall, the topography plots show a strong correlation with the VRAD map, but importantly also display a large amount of variability across time.

Outlook

Where next for M/EEG Source Reconstruction Research?

Let us begin by limiting our discussion to MAP solutions to the inverse problem. We restate Equation (2.10) here for the convenience of the reader:

$$\hat{\mathbf{X}} = \mathbf{C}_X\mathbf{H}^T[\mathbf{C}_n + \mathbf{H}\mathbf{C}_X\mathbf{H}^T]^{-1}\mathbf{Y},$$

where any dynamics in $\mathbf{C}_X$ and $\mathbf{C}_n$ are implicit. Let us consider each of these elements in turn:

- $\mathbf{C_X}$ - the source space covariance matrix. Now that we have introduced dynamics to the priors, it is hard to know what advances can be made on an isolated single subject basis. We may be able to leverage cross-subject information by learning common modes of activation on the population level, and employing some form of hierarchical temporal modelling in the future.
- **H** - the lead field matrix. This can be improved on both engineering and modelling fronts. Practically, there is always scope for sensors to be improved in terms of sensitivity and noise floor during the manufacturing process such

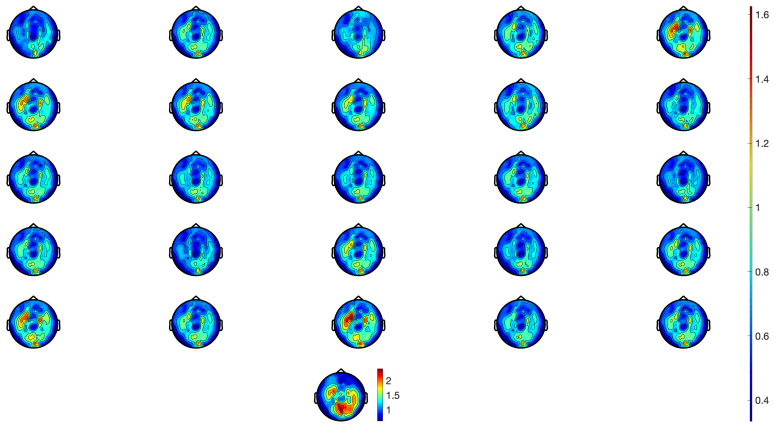


Figure 7.2: Non-categorical temporal modelling captures more temporal diversity in the variance. The instantaneous VRAD variances (rows 1-5) are shown at 25 random timepoints from when state 19 of the HMM was visited (shown below). Notice how, across all timepoints, the maps are similar in topography and scale between the two techniques, yet also how there is substantial timepoint to timepoint variability in the VRAD approach. This is particularly true in the occasional recruitment of a dipole over right frontal areas (not seen in every map). The HMM is not able to capture such temporal variability without constructing more states.

that derived matrices align closely with the physical reality of the system. OPM arrays will in the future be customisable to increase the SNR of a chosen brain region, thus changing the values of this matrix.

From a modelling perspective, there is ongoing research into improving the veracity of computational approaches of forward models. One of the current drawbacks of OPMs is their requirement to be placed in a low (ideally, $0nT$) background field. When this criterion is not met, so-called “gain errors" are introduced into the sensors, such that the theoretical lead field does not tally with the recorded signals (Duque-Muñoz et al., 2019). Computational techniques, perhaps modelling the lead field matrix as an initial estimate, $\mathbf{H}_0$, plus some data driven corrections, $\mathbf{H} = \mathbf{H}_0 + \sum_i c_i \mathbf{H}_i$ may help to solve this issue.

- $\mathbf{C}_n$ - the noise covariance matrix. We briefly entertained the idea of using time-varying noise regularisation in this thesis, and indeed, did so *implicitly* when using TABFER (c.f. the cardiac components). As stressed in the theory section, this matrix is particularly important for tomographic approaches where there is only one source prior (MNE, LORETA, etc.). We could easily expand the regularisation term to incorporate more than one dynamic noise prior, perhaps learnt from empty room recordings.

Future applications of TASER

We showed in this thesis that TASER can give more accurate source space images in the context of a precision MEG dataset. This could, one day, mean that TASER would be a suitable inversion algorithm for clinical work, where precision and accuracy are a requirement for surgical intervention (Rampp et al., 2019). Equally, TASER could be used for small datasets where source images of individual subjects are likely to be highly interrogated (i.e. personalised medicine). Our simulation and

experimental analysis showed that using $N_W > 1$ can return different timeseries. This makes functional connectivity studies a natural further avenue of research. Simulations show that deep (e.g. hippocampal regions) stand to benefit from the use of TASER in the future. As outlined earlier in this chapter, the only way of improving source reconstruction in a naturalistic or resting state paradigm is to take an unsupervised approach to the learning of weights. All of these future avenues will perhaps best be realised within a generative-model based framework where formal model comparisons can be made.

Conclusion

We began this thesis by acknowledging that the brain was a dynamical system, whose distribution of activity varies considerably from one instance to the next. MEG is unique in its ability to provide 5 dimensional images of this evolution of neural processes, once the ill-posed inverse problem is tackled. Our work has shown that failing to account for changes in the spatial layout of power can degrade images of brain activity, and that unsupervised learning approaches can go a long way to better elucidate what is really going on inside the skull.

A

Appendix A: Coregistration

Consider (at least) three points in $\mathbb{R}^3$ in two different spaces (e.g. MEG and MRI). These points are denoted as $\{x, y, z\}$ and $\{x', y', z'\}$, respectively. Our aim is to simply learn a rigid body transformation between the two systems, $\mathbf{M}_1$, such that we can place a detailed description of the subject's brain (the MRI) inside the array of sensors:

$$\begin{bmatrix} x' \\ y' \\ z' \\ 1 \end{bmatrix} = \mathbf{M}_1 \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}. \tag{A.1}$$

Making the use of homogenous coordinates[1], $\mathbf{M}_1$ can be decomposed as

$$\mathbf{M}_1 = \begin{bmatrix} 1 & 0 & 0 & \Delta x \\ 0 & 1 & 0 & \Delta y \\ 0 & 0 & 1 & \Delta z \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & cos(\theta_x) & -sin(\theta_x) & 0 \\ 0 & sin(\theta_x) & cos(\theta_x) & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \tag{A.2}$$

The first matrix encodes translations in the $x/y/z$ directions and the second matrix represents rotations around the x axis (for the purposes of our discussion we ignore the other two directions). We can frame the learning of this matrix as a least squares problem and find the rotation matrix, $\mathbf{R}$, by use of a singular value decomposition on the correlation matrix $\mathbf{H}$:

$$\mathbf{H} = \mathbf{S}^T\mathbf{M} \tag{A.3}$$

$$\mathbf{H} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T \tag{A.4}$$

$$\mathbf{R} = \mathbf{V}\mathbf{U}^T \tag{A.5}$$

where $\mathbf{M}$ and $\mathbf{S}$ are the (zero-centred) set of points ($\mathbb{R}^{N\times 3}$) in the target and initial space, respectively. This was first shown by Schönemann (1966). The translation vector can be found simply via

$$\mathbf{t} = \mathbf{m} - \mathbf{R}\mathbf{s} \tag{A.6}$$

where $\mathbf{m}$ and $\mathbf{s}$ are the centroids of the original set of points in the two spaces. The final translation matrix $\mathbf{M}_1$ is hence given by the augmentation of these two results.

It should be noted that the least squares nature of the cost function makes the

[1] All this means is that we add a fourth dimension to the problem, $[x, y, z] \rightarrow [x, y, z, w]$ where typically $w := 1$. This fourth dimension is disregarded after the transformation matrix has been learnt and makes the maths more convenient. Formally, we have placed the 3D coordinates on a hyperplane in 4 dimensions.

coregistration process particularly sensitive to outlier data points. Practically, this issue often manifests itself when many headshape points are acquired alongside the three “traditional" fiducial points.

B

Appendix B: MEG and EEG Forward Models

The M/EEG inverse problem is intimidating, cumbersome and dependent upon a lot of user choices. In some respects, the forward problem can be considered a tamer beast (so long as one can stomach the vector calculus[1] and has faith in their estimates for different tissue conductivities[2]) as it is fully determined by well-defined physical formulae. In this section, we will explore the MEG forward problem and outline how one goes about calculating it in a simplified case.

It should be stressed that none of the details in this section are novel, yet no thesis on source localisation would be complete without entertaining Maxwell's Equations.

Problem Statement

Simply put, the forward problem answers "what are the electro/magnetic fields that one would expect to measure if a known current source was placed in a conductive medium?"

Formulation

There are different ways of arriving at a solution to the M/EEG forward problem (Sekihara and Nagarajan, 2015; Sarvas, 1987), but invariably all derivations begin by considering the non-relativistic form of Maxwell's Equations. Consider a volume, G, enclosed by a surface, S. The electromagnetic fields originating from such a conductor are governed by:

$$\nabla.\mathbf{E} = \frac{\rho}{\epsilon} \tag{B.1}$$

$$\nabla.\mathbf{B} = 0 \tag{B.2}$$

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \tag{B.3}$$

$$\nabla \times \mathbf{B} = \mu\mathbf{J} + \mu\epsilon\frac{\partial \mathbf{E}}{\partial t} \tag{B.4}$$

[1] This does not necessarily include the author!

[2] There is evidence suggesting that gaining this confidence may be tricky (McCann et al., 2019).

where $\mathbf{E}$ is the electric field, $\mathbf{B}$ is the magnetic field, ρ is the charge density, μ is the magnetic permeability and ϵ is the permittivity of the conductor.

The convention within the field of bioelectromagnetism is to now simplify these expressions using the so-called “quasistatic approximation" to Maxwell’s Equations. The lemma is that because the fields we interest ourselves in are typically of low frequency ($0.5-150Hz$) with a $1/f$ profile, the time derivatives in Maxwell’s Equations can be neglected (Geselowitz, 1967; Hämäläinen et al., 1993; Plonsey, 1982). Indeed, analysis by Sekihara and Nagarajan (2015) shows that errors due to capacitive effects are on the order of 1% for oscillations at $100Hz$ and only 3% at $1kHz$[3]. Moreover, as our signals of interest (the magnetic fields) propagate only a short distance (circa $10-20cm$ in a cryogenic MEG system, and about half that for an OPM system) at the speed of light, one can also neglect any effects due to propagation times. To summarise, the temporal derivatives above can and are approximated to be zero.

Above, $\mathbf{J}$ is the total current density, which we model as the linear summation of primary and secondary currents:

$$\mathbf{J} = \mathbf{J}^P + \mathbf{J}^S, \tag{B.5}$$

$$= \mathbf{J}^P - \sigma\nabla V, \tag{B.6}$$

where $\mathbf{J}^P$ are the primary currents, that is the flow of charges in neurons, and $\mathbf{J}^S$ are the so-called “secondary" or Ohmic extracellular currents which arise as a biproduct of the primary currents. Crucially, σ is the local conductivity of the volume, $\sigma = \sigma(\mathbf{r}')$. We shall return to this important term shortly.

Now that we have defined our problem and neural generators, we can begin to derive a solution. A very thorough and formal set of derivations built upon potential functions and gauge transformations are provided in Sekihara and Nagarajan (2015), where the magnetic forward problem is considered in the cases of bounded and unbounded and homogeneous and piecewise-constant conductive mediums. Whilst these mathematical formulisms are both elegant and important (especially for the EEG source modeller), we shall not consider such methods in any great detail here. However, we will take two key findings from this book.

The first is that by adopting the formal potential approach, we eventually arrive at the Ampère-Laplace law:

$$\mathbf{B}(\mathbf{r}) = \frac{\mu_0}{4\pi}\int_G \mathbf{J}(\mathbf{r}') \times \frac{\mathbf{r}-\mathbf{r}'}{|(\mathbf{r}-\mathbf{r}')|^3} dv', \tag{B.7}$$

where $\mathbf{r}'$ is the location of a dipole inside the head, $\mathbf{r}$ represents a point outside of the head, and dv' is a volume element within the head. We have made the substitution $\mu = \mu_0$, the permeability of free space.

The second teaching is that when we do not make the (arguably crude) assumption that the head is a homogenous spherical conductor, we cannot neglect the effects of spatially varying conductivity values. Let us show this. Consider the realistic case that the head can be modelled as a bounded conductor with a piecewise constant conductivity. Making the substitution $\mathbf{R} = \mathbf{r} - \mathbf{r}'$ and using our definition of the total current density, we may write

$$\mathbf{B}(\mathbf{r}) = \frac{\mu_0}{4\pi}\int_G \left(\mathbf{J}^P - \sigma(\mathbf{r}')\nabla V(\mathbf{r}')\right) \times \frac{\mathbf{R}}{|\mathbf{R}|^3} dv'. \tag{B.8}$$

To ease the notation, we will here emphasise the fact here that $\mathbf{J}^P = \mathbf{J}^P(\mathbf{r}')$ and $V = V(\mathbf{r}')$ and no longer explicitly label this spatial dependency. Equation (B.8)

[3] Naturally one could therefore envisage such modelling errors becoming relevant in ultra-high frequency or epilepsy studies.

has significant implications in our efforts to calculate a MEG forward model. In words, to solve the forward problem in the most realistic and broadest setting for MEG, we must a) estimate the local conductivity profile of the head (an arduous task, which is still a very active area of research, especially as it is believed that conductivity profiles exhibit significant inter-subject variability, see Liu et al. (2018), for example), and b) solve the EEG forward problem, such that we have a good estimate for the local electric potentials created by the primary and secondary currents.

Proceeding, using vector calculus identities and the generalised Stokes' theorem, eventually one arrives at the generalised Geselowitz formula (Geselowitz, 1970):

$$\mathbf{B}(\mathbf{r}) = \mathbf{B}_0(\mathbf{r}) - \frac{\mu_0}{4\pi}\sum_{j=1}^{\omega}\left[\sigma_j - \sigma_j'\right]\int_{S_j} V\hat{\mathbf{n}}(\mathbf{r}') \times \mathbf{R}dS \tag{B.9}$$

where $\mathbf{B}_0(\mathbf{r})$ is the magnetic field arising from the primary currents in an infinitely sized homogeneous conductor, S_j is a surface within the volume G, $\hat{\mathbf{n}}(\mathbf{r}')$ is a normal vector to each of the surfaces and there are a total of ω regions inside the volume. Here, we have modelled the head as a heterogeneous conductor with different local conductivities, σ_j.

Evaluating the second term in Equation (B.9) for an arbitrary source space geometry calls for numerical methods. The more pragmatic MEG physicist typically now introduces the assumption that the head is a spherical conductor, such that the radial component of the magnetic field is given by

$$B_r(\mathbf{r}) = \mathbf{B}(\mathbf{r}).\mathbf{e}_r \tag{B.10}$$

$$= \mathbf{B}_0(\mathbf{r}).\mathbf{e}_r - \frac{\mu_0}{4\pi}\sum_{j=1}^{\omega}\left[\sigma_j - \sigma_j'\right]\int_{S_j} V\hat{\mathbf{n}}(\mathbf{r}') \times \mathbf{R}.\mathbf{e}_r dS \tag{B.11}$$

where $\mathbf{e}_r$ is a radial unit vector. By the definition of the cross and subsequent dot products, this results in all of the terms within the surface integrand equating to zero (as $\hat{\mathbf{n}}(\mathbf{r}') = \mathbf{e}_r$) and so we only have to consider the primary currents when measuring the radial component of the magnetic field. Conveniently, this is not a function of the conductivity of the volume.

Recall that the radial component of the magnetic field due to primary currents is given by

$$B_r(\mathbf{r}) = \frac{\mu_0}{4\pi}\int_G \left(\mathbf{J}^P \times \frac{\mathbf{R}}{|\mathbf{R}|^3}\right).\mathbf{e}_r dv'. \tag{B.12}$$

Now, let us assume that currents inside the head are point sources. This lets us model the sources of magnetic fields with a Dirac delta function:

$$\mathbf{J}^P = \mathbf{Q}\delta\left(\mathbf{r}' - \mathbf{r}_Q\right), \tag{B.13}$$

where $\mathbf{Q}$ is a current dipole and

$$\delta\left(\mathbf{r}' - \mathbf{r}_Q\right) = \begin{cases} 0 & \text{if } \mathbf{r}' = \mathbf{r}_Q \\ 1 & \text{otherwise} \end{cases}. \tag{B.14}$$

This assumption greatly simplifies the above integral. It is trivial to see now that

$$B_r(\mathbf{r}) = \frac{\mu_0}{4\pi}\mathbf{Q}(\mathbf{r}_Q) \times \frac{\mathbf{r} - \mathbf{r}_Q}{|\mathbf{r} - \mathbf{r}_Q|^3}.\mathbf{e}_r. \tag{B.15}$$

The picture is complete once we take the other components of the magnetic field into account. Starting with the magnetic potential, $U(\mathbf{r})$, after assuming that a

dipole exists at the location $\mathbf{r}_Q$ and integrating the scalar potential along a line integral between the origin and infinity, we arrive at an expression for the potential

$$U(\mathbf{r}) = -\frac{1}{4\pi}\frac{\mathbf{Q}\times\mathbf{r}_0.\mathbf{r}}{\Lambda}, \tag{B.16}$$

where

$$\Lambda = |\mathbf{r}-\mathbf{r}_Q|\left(|\mathbf{r}-\mathbf{r}_Q|\mathbf{r}+|\mathbf{r}|^2-\mathbf{r}_Q.\mathbf{r}\right). \tag{B.17}$$

Armed with this expression, the Sarvas formula for a homogeneous spherical conductor is finally given by

$$B(\mathbf{r}) = -\mu_0\nabla U(\mathbf{r}) \tag{B.18}$$

$$= \frac{\mu_0}{4\pi\Lambda^2}\left[\Lambda\mathbf{Q}\times\mathbf{r}_0-\{(\mathbf{Q}\times\mathbf{r}_0).\mathbf{r}\}\nabla\Lambda\right]. \tag{B.19}$$

C

Appendix C: Deriving Beamformer Weights

We use the same notation here as in the main body of this thesis. The sensor level recordings are written as $\mathbf{Y}$, $\mathbf{C}_Y$ is its covariance, $\mathbf{h}$ is a column of the lead field matrix and $\mathbf{w}$ are the beamformer weights.

We now show how the expression for the LCMV beamformer weights can be derived. We start by writing down our goal to minimise the power of the solution, alongside the constraint that a unit source at each location, i, has a unit gain:

$$\mathbf{w}_i = \min_{\mathbf{w}_i} \mathbf{w}_i^T \mathbf{C}_Y \mathbf{w}_i \quad \text{subject to} \quad \mathbf{w}_i^T \mathbf{h}_i = 1. \tag{C.1}$$

A solution is found to the above minimisation problem by means of introducing a Lagrange multiplier, λ. The Lagrangian function then reads

$$\mathcal{L}(\mathbf{w}, \lambda) = \mathbf{w}_i^T \mathbf{C}_Y \mathbf{w}_i + \lambda \left(\mathbf{w}_i^T \mathbf{h}_i - 1 \right) \tag{C.2}$$

which is the function we now seek to minimise with respect to $\mathbf{w}$. Differentiating Equation (C.2) and equating the result to zero,

$$\frac{\partial \mathcal{L}(\mathbf{w}, \lambda)}{\partial \mathbf{w}_i} = 2\mathbf{C}_Y \mathbf{w}_i + \lambda \mathbf{h}_i = 0, \tag{C.3}$$

where we used the general result that $\partial \mathbf{x}^T \mathbf{A} \mathbf{x} / \partial \mathbf{x} = \left(\mathbf{A} + \mathbf{A}^T \right) \mathbf{x}$, we find that the beamformer weights are given by

$$\mathbf{w}_i = -\lambda \mathbf{C}_Y^{-1} \mathbf{h}_i / 2. \tag{C.4}$$

To eliminate λ from Equation (C.4) we substitute this expression into the original constraint introduced at the start of the derivation. Doing this, we obtain

$$\lambda = -\frac{2}{\mathbf{h}_i^T \mathbf{C}_Y^{-1} \mathbf{h}_i}. \tag{C.5}$$

Plugging this expression into Equation (C.4), we obtain the expression for the LCMV beamformer weights:

$$\mathbf{w}_{BF,i} = \frac{\mathbf{C}_Y^{-1} \mathbf{h}_i}{\mathbf{h}_i^T \mathbf{C}_Y^{-1} \mathbf{h}_i}. \tag{C.6}$$

D

Appendix D: MAP Equation Derivation

For a given model, $\mathcal{M}$, Bayes' theorem tells us that the posterior distribution for the source space currents $\mathbf{X}$ is given by

$$p(\mathbf{X}|\mathbf{Y}) = \frac{p(\mathbf{Y}|\mathbf{X})p(\mathbf{X})}{p(\mathbf{Y})}, \tag{D.1}$$

where $\mathbf{Y}$ are the observable data. Evaluating $p(\mathbf{Y}) = \int p(\mathbf{Y}|\mathbf{X})p(\mathbf{X})d\mathbf{X}$ is often intractable. If we are only to concern ourselves with point estimates for $\hat{\mathbf{X}}$, then we may neglect this troublesome denominator and instead only concern ourselves with the values which maximise the posterior. This is known as the *maximum a posteriori* (MAP) solution:

$$p(\mathbf{X}|\mathbf{Y}) \propto p(\mathbf{Y}|\mathbf{X})p(\mathbf{X}) \tag{D.2}$$

The putative sources are modelled as zero-mean normal distributions with covariance $\mathbf{C}_X$, whilst the sensor level measurements are a product of the lead fields and the neuronal activity, with some additive zero-mean normally distributed noise of covariance $\mathbf{C_n}$:

$$p(\mathbf{X}) = N(\mathbf{X}; \mathbf{0}, \mathbf{C}_X) = m \exp\left[-\frac{1}{2}\mathbf{X}^T\mathbf{C}_X^{-1}\mathbf{X}\right] \tag{D.3}$$

$$p(\mathbf{Y}|\mathbf{X}) = N(\mathbf{Y}; \mathbf{HX}, \mathbf{C}_n) = k \exp\left[-\frac{1}{2}(\mathbf{HX}-\mathbf{Y})^T\mathbf{C}_n^{-1}(\mathbf{HX}-\mathbf{Y})\right] \tag{D.4}$$

where m and k are constants. Hence

$$p(\mathbf{X}|\mathbf{Y})p(\mathbf{X}) \propto \exp\left[-\frac{1}{2}(\mathbf{HX}-\mathbf{Y})^T\mathbf{C}_n^{-1}(\mathbf{HX}-\mathbf{Y}) - \frac{1}{2}\mathbf{X}^T\mathbf{C}_X^{-1}\mathbf{X}\right]. \tag{D.5}$$

We now need to maximise Equation (D.5). Taking logs, differentiating with respect to the source space currents and equating this to zero we find that

$$\frac{d(log(p(\mathbf{X}|\mathbf{Y})p(\mathbf{X})))}{d\mathbf{X}} = 0 = -\mathbf{H}^T\mathbf{C}_n^{-1}(\mathbf{HX}-\mathbf{Y}) - \mathbf{C}_X^{-1}\mathbf{X}, \tag{D.6}$$

and hence the MAP solution is given by

$$\hat{\mathbf{X}} = \mathbf{C}_X\mathbf{H}^T[\,\mathbf{C}_n + \mathbf{H}\mathbf{C}_X\mathbf{H}^T]^{-1}\mathbf{Y} \tag{D.7}$$

That’s the end, and that’s the start of it

Bibliography

Seppo P Ahlfors, Jooman Han, John W Belliveau, and Matti S Hämäläinen. Sensitivity of meg and eeg to source orientation. *Brain topography*, 23(3):227–232, 2010.

Gregory E Alberto, Jennifer R Stapleton-Kotloski, David J Klorig, Emily R Rogers, Christos E Constantinidis, James Daunais, and Dwayne W Godwin. Optogenetic activation of nonhuman primate cortical and subcortical brain circuits highlights detection capabilities of meg source imaging. *bioRxiv*, 2020.

Ann Elisabeth Albright Blomberg, Andreas Austeng, Roy Edgar Hansen, and Stig Asle Vaksvik Synnes. Improving sonar performance in shallow water using adaptive beamforming. *IEEE Journal of Oceanic Engineering*, 38(2):297–307, 2013. doi: 10.1109/JOE.2012.2226643.

John Ashburner and Karl J Friston. Unified segmentation. *Neuroimage*, 26(3):839–851, 2005.

Yohan Attal and Denis Schwartz. Assessment of subcortical source localization using deep brain activity imaging model with minimum norm operators: a meg study. *Plos one*, 8(3):e59856, 2013.

Jimmy Lei Ba, Jamie Ryan Kiros, and Geoffrey E Hinton. Layer normalization. *arXiv preprint arXiv:1607.06450*, 2016.

George Backus and Freeman Gilbert. Uniqueness in the inversion of inaccurate gross earth data. *Philosophical Transactions of the Royal Society of London. Series A, Mathematical and Physical Sciences*, 266(1173):123–192, 1970.

Adam P Baker, Matthew J Brookes, Iead A Rezek, Stephen M Smith, Timothy Behrens, Penny J Probert Smith, and Mark Woolrich. Fast transient networks in spontaneous human brain activity. *Elife*, 3:e01867, 2014.

Pierre Baldi, Søren Brunak, and Francis Bach. *Bioinformatics: the machine learning approach*. MIT press, 2001.

Siqi Bao and Albert CS Chung. Multi-scale structured cnn with label consistency for brain mr image segmentation. *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization*, 6(1):113–117, 2018.

John Bardeen, Leon N Cooper, and John Robert Schrieffer. Theory of superconductivity. *Physical review*, 108(5):1175, 1957.

Eleanor L Barratt, Susan T Francis, Peter G Morris, and Matthew J Brookes. Mapping the topological organisation of beta oscillations in motor cortex using meg. *NeuroImage*, 181:831–844, 2018.

Justin Bayer and Christian Osendorfer. Learning stochastic recurrent networks. *arXiv preprint arXiv:1411.7610*, 2014.

Christian F Beckmann, Marilena DeLuca, Joseph T Devlin, and Stephen M Smith. Investigations into resting-state connectivity using independent component analysis. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 360(1457):1001–1013, 2005.

John M Bekkers. Pyramidal neurons. *Current biology*, 21(24):R975, 2011.

Paolo Belardinelli, Erick Ortiz, Gareth Barnes, Uta Noppeney, and Hubert Preissl. Source reconstruction accuracy of meg and eeg bayesian inversion approaches. *PloS one*, 7(12):e51985, 2012.

Shlomo Bentin, Truett Allison, Aina Puce, Erik Perez, and Gregory McCarthy. Electrophysiological studies of face perception in humans. *Journal of cognitive neuroscience*, 8(6):551–565, 1996.

Hans Berger. Über das elektroenkephalogramm des menschen. *Archiv für psychiatrie und nervenkrankheiten*, 87(1):527–570, 1929.

Christopher M Bishop. Pattern recognition. *Machine learning*, 128(9), 2006.

Felix Bloch. Nuclear induction. *Physical review*, 70(7-8):460, 1946.

Alexander A Borbély, Fritz Baumann, Daniel Brandeis, Inge Strauch, and Dietrich Lehmann. Sleep deprivation: effect on sleep stages and eeg power density in man. *Electroencephalography and clinical neurophysiology*, 51(5):483–493, 1981.

Jelmer P Borst and John R Anderson. The discovery of processing stages: Analyzing eeg data with hidden semi-markov models. *NeuroImage*, 108:60–73, 2015.

Elena Boto, Niall Holmes, James Leggett, Gillian Roberts, Vishal Shah, Sofie S Meyer, Leonardo Duque Muñoz, Karen J Mullinger, Tim M Tierney, Sven Bestmann, et al. Moving magnetoencephalography towards real-world applications with a wearable system. *Nature*, 555(7698):657–661, 2018.

Elena Boto, Ryan M Hill, Molly Rea, Niall Holmes, Zelekha A Seedat, James Leggett, Vishal Shah, James Osborne, Richard Bowtell, and Matthew J Brookes. Measuring functional connectivity with wearable meg. *NeuroImage*, 230:117815, 2021.

Samuel R Bowman, Luke Vilnis, Oriol Vinyals, Andrew M Dai, Rafal Jozefowicz, and Samy Bengio. Generating sentences from a continuous space. *arXiv preprint arXiv:1511.06349*, 2015.

Korbinian Brodmann. *Vergleichende Lokalisationslehre der Grosshirnrinde in ihren Prinzipien dargestellt auf Grund des Zellenbaues*. Barth, 1909.

Matthew J Brookes, Andrew M Gibson, Stephen D Hall, Paul L Furlong, Gareth R Barnes, Arjan Hillebrand, Krish D Singh, Ian E Holliday, Sue T Francis, and Peter G Morris. A general linear model for meg beamformer imaging. *NeuroImage*, 23(3): 936–946, 2004.

Matthew J Brookes, Claire M Stevenson, Gareth R Barnes, Arjan Hillebrand, Michael IG Simpson, Susan T Francis, and Peter G Morris. Beamformer reconstruction of correlated sources using a modified source model. *Neuroimage*, 34(4):1454–1465, 2007.

Matthew J Brookes, Jiri Vrba, Stephen E Robinson, Claire M Stevenson, Andrew M Peters, Gareth R Barnes, Arjan Hillebrand, and Peter G Morris. Optimising experimental design for meg beamformer imaging. *Neuroimage*, 39(4):1788–1802, 2008.

Matthew J Brookes, Johanna M Zumer, Claire M Stevenson, Joanne R Hale, Gareth R Barnes, Jiri Vrba, and Peter G Morris. Investigating spatial specificity and data averaging in meg. *Neuroimage*, 49(1):525–538, 2010.

Matthew J Brookes, Mark Woolrich, Henry Luckhoo, Darren Price, Joanne R Hale, Mary C Stephenson, Gareth R Barnes, Stephen M Smith, and Peter G Morris. Investigating the electrophysiological basis of resting state networks using magnetoencephalography. *Proceedings of the National Academy of Sciences*, 108 (40):16783–16788, 2011.

Joana Cabral, Diego Vidaurre, Paulo Marques, Ricardo Magalhães, Pedro Silva Moreira, José Miguel Soares, Gustavo Deco, Nuno Sousa, and Morten L Kringelbach. Cognitive performance in healthy older adults relates to spontaneous switching between states of functional connectivity during rest. *Scientific reports*, 7(1):1–13, 2017.

Stanislas Chambon, Mathieu N Galtier, Pierrick J Arnal, Gilles Wainrib, and Alexandre Gramfort. A deep learning architecture for temporal sleep stage classification using multivariate and multimodal time series. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 26(4):758–769, 2018.

Omar Chehab, Alexandre Defossez, Jean-Christophe Loiseau, Alexandre Gramfort, and Jean-Remi King. Deep recurrent encoder: A scalable end-to-end network to model brain signals. *arXiv preprint arXiv:2103.02339*, 2021.

Liang-Chieh Chen, George Papandreou, Iasonas Kokkinos, Kevin Murphy, and Alan L Yuille. Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs. *IEEE transactions on pattern analysis and machine intelligence*, 40(4):834–848, 2018.

Douglas Cheyne, Andreea C Bostan, William Gaetz, and Elizabeth W Pang. Event-related beamforming: a robust method for presurgical functional mapping using meg. *Clinical Neurophysiology*, 118(8):1691–1704, 2007.

Douglas Owen Cheyne. Meg studies of sensorimotor rhythms: a review. *Experimental neurology*, 245:27–39, 2013.

David Cohen. Magnetoencephalography: evidence of magnetic fields produced by alpha-rhythm currents. *Science*, 161(3843):784–786, 1968.

Giles L Colclough, Mark W Woolrich, PK Tewarie, Matthew J Brookes, Andrew J Quinn, and Stephen M Smith. How reliable are meg resting-state connectivity metrics? *Neuroimage*, 138:284–293, 2016.

Michael W Cole, Danielle S Bassett, Jonathan D Power, Todd S Braver, and Steven E Petersen. Intrinsic and task-evoked network architectures of the human brain. *Neuron*, 83(1):238–251, 2014.

Christine M Constantinople and Randy M Bruno. Deep cortical layers are activated directly by thalamus. *Science*, 340(6140):1591–1594, 2013.

Brian R Cornwell, Linda L Johnson, Tom Holroyd, Frederick W Carver, and Christian Grillon. Human hippocampal and parahippocampal theta during goal-directed spatial navigation predicts performance on a virtual morris water maze. *Journal of Neuroscience*, 28(23):5983–5990, 2008.

Sarang S Dalal, Adrian G Guggisberg, Erik Edwards, Kensuke Sekihara, Anne M Findlay, Ryan T Canolty, Mitchel S Berger, Robert T Knight, Nicholas M Barbaro, Heidi E Kirsch, et al. Five-dimensional neuroimaging: localization of the time–frequency dynamics of cortical activity. *Neuroimage*, 40(4):1686–1700, 2008.

Anders M Dale, Arthur K Liu, Bruce R Fischl, Randy L Buckner, John W Belliveau, Jeffrey D Lewine, and Eric Halgren. Dynamic statistical parametric mapping: combining fmri and meg for high-resolution imaging of cortical activity. *Neuron*, 26(1):55–67, 2000.

Francesco De Pasquale, Stefania Della Penna, Abraham Z Snyder, Christopher Lewis, Dante Mantini, Laura Marzetti, Paolo Belardinelli, Luca Ciancetta, Vittorio Pizzella, Gian Luca Romani, et al. Temporal dynamics of spontaneous meg activity in brain networks. *Proceedings of the National Academy of Sciences*, 107(13):6040–6045, 2010.

Rahul S Desikan, Florent Ségonne, Bruce Fischl, Brian T Quinn, Bradford C Dickerson, Deborah Blacker, Randy L Buckner, Anders M Dale, R Paul Maguire, Bradley T Hyman, et al. An automated labeling system for subdividing the human cerebral cortex on mri scans into gyral based regions of interest. *Neuroimage*, 31(3):968–980, 2006.

Stavros I Dimitriadis, Bethany Routley, David E Linden, and Krish D Singh. Reliability of static and dynamic network metrics in the resting-state: a meg-beamformed connectivity analysis. *Frontiers in neuroscience*, 12:506, 2018.

Mithun Diwakar, Ming-Xiong Huang, Ramesh Srinivasan, Deborah L Harrington, Ashley Robb, Annemarie Angeles, Laura Muzzatti, Reza Pakdaman, Tao Song, Rebecca J Theilmann, et al. Dual-core beamformer for obtaining highly correlated neuronal networks in meg. *NeuroImage*, 54(1):253–263, 2011.

Leonardo Duque-Muñoz, Tim M Tierney, Sofie S Meyer, Elena Boto, Niall Holmes, Gillian Roberts, James Leggett, JF Vargas-Bonilla, Richard Bowtell, Matthew J Brookes, et al. Data-driven model optimization for optically pumped magnetometer sensor arrays. *Human Brain Mapping*, 40(15):4357–4369, 2019.

Henri M Duvernoy. *The human hippocampus: functional anatomy, vascularization and serial sections with MRI*. Springer Science & Business Media, 2005.

Denis A Engemann and Alexandre Gramfort. Automated model selection in covariance estimation and spatial whitening of meg and eeg signals. *NeuroImage*, 108:328–342, 2015.

Joseph Feingold, Daniel J Gibson, Brian DePasquale, and Ann M Graybiel. Bursts of beta oscillation differentiate postperformance activity in the striatum and motor cortex of monkeys performing movement tasks. *Proceedings of the National Academy of Sciences*, 112(44):13687–13692, 2015.

Bruce Fischl. Freesurfer. *Neuroimage*, 62(2):774–781, 2012.

Alison E Fisher, Gareth R Barnes, Arjan Hillebrand, Caroline Burrow, Paul L Furlong, and Ian E Holliday. Can you tell your clunis from your cubitus? a benchmark for functional imaging. *Bmj*, 329(7480):1492–1493, 2004.

Meire Fortunato, Charles Blundell, and Oriol Vinyals. Bayesian recurrent neural networks. *arXiv preprint arXiv:1704.02798*, 2017.

Karl Friston, Jérémie Mattout, Nelson Trujillo-Barreto, John Ashburner, and Will Penny. Variational free energy and the laplace approximation. *Neuroimage*, 34(1):220–234, 2007.

Karl Friston, Lee Harrison, Jean Daunizeau, Stefan Kiebel, Christophe Phillips, Nelson Trujillo-Barreto, Richard Henson, Guillaume Flandin, and Jérémie Mattout. Multiple sparse priors for the m/eeg inverse problem. *NeuroImage*, 39(3):1104–1120, 2008.

Karl J Friston, William Penny, Christophe Phillips, S Kiebel, G Hinton, and John Ashburner. Classical and bayesian inference in neuroimaging: theory. *NeuroImage*, 16(2):465–483, 2002.

Adam Fry, Karen J Mullinger, George C O'Neill, Eleanor L Barratt, Peter G Morris, Markus Bauer, Jonathan P Folland, and Matthew J Brookes. Modulation of post-movement beta rebound by contraction force and rate of force development. *Human brain mapping*, 37(7):2493–2511, 2016.

Aurelien Geron. *Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow: Concepts, tools, and techniques to build intelligent systems*. O'Reilly Media, 2019.

D Geselowitz. On the magnetic field generated outside an inhomogeneous volume conductor by internal current sources. *IEEE Transactions on Magnetics*, 6(2):346–347, 1970.

David B Geselowitz. On bioelectric potentials in an inhomogeneous volume conductor. *Biophysical journal*, 7(1):1–11, 1967.

Zoubin Ghahramani. An introduction to hidden markov models and bayesian networks. In *Hidden Markov models: applications in computer vision*, pages 9–41. World Scientific, 2001.

D. E. Goldman. Potential, impedance, and rectification in membranes. *J Gen Physiol*, 27(1):37–60, 1943. URL http://www.ncbi.nlm.nih.gov/pubmed/19873371.

Irina F Gorodnitsky, John S George, and Bhaskar D Rao. Neuromagnetic source imaging with focuss: a recursive weighted minimum norm algorithm. *Electroencephalography and clinical Neurophysiology*, 95(4):231–251, 1995a.

Irina F Gorodnitsky, John S George, and Bhaskar D Rao. Neuromagnetic source imaging with focuss: a recursive weighted minimum norm algorithm. *Electroencephalography and clinical Neurophysiology*, 95(4):231–251, 1995b.

Alex Graves and Jürgen Schmidhuber. Offline handwriting recognition with multidimensional recurrent neural networks. In *Advances in neural information processing systems*, pages 545–552, 2009.

Karol Gregor, Ivo Danihelka, Alex Graves, Danilo Jimenez Rezende, and Daan Wierstra. Draw: A recurrent neural network for image generation. *arXiv preprint arXiv:1502.04623*, 2015.

Joachim Groß, Lars Timmermann, Jan Kujala, Martin Dirks, Frank Schmitz, R Salmelin, and Alfons Schnitzler. The neural basis of intermittent motor control in humans. *Proceedings of the National Academy of Sciences*, 99(4):2299–2302, 2002.

Joachim Gross, Sylvain Baillet, Gareth R Barnes, Richard N Henson, Arjan Hillebrand, Ole Jensen, Karim Jerbi, Vladimir Litvak, Burkhard Maess, Robert Oostenveld, et al. Good practice for conducting and reporting meg research. *Neuroimage*, 65:349–363, 2013.

Michael BH Hall, Ida A Nissen, Elisabeth CW van Straaten, Paul L Furlong, Caroline Witton, Elaine Foley, Stefano Seri, and Arjan Hillebrand. An evaluation of kurtosis beamforming in magnetoencephalography to localize the epileptogenic zone in drug resistant epilepsy patients. *Clinical Neurophysiology*, 129(6):1221–1229, 2018.

Matti Hämäläinen, Riitta Hari, Risto J Ilmoniemi, Jukka Knuutila, and Olli V Lounasmaa. Magnetoencephalography - theory, instrumentation, and applications to noninvasive studies of the working human brain. *Reviews of modern Physics*, 65(2):413, 1993.

Matti S Hämäläinen and Risto J Ilmoniemi. Interpreting magnetic fields of the brain: minimum norm estimates. *Medical & biological engineering & computing*, 32(1): 35–42, 1994.

Matti S Hamalainen and Jukka Sarvas. Realistic conductivity geometry model of the human head for interpretation of neuromagnetic data. *IEEE transactions on biomedical engineering*, 36(2):165–171, 1989.

R Hari, K Aittoniemi, M-L Järvinen, T Katila, and T Varpula. Auditory evoked transient and sustained magnetic fields of the human brain localization of neural generators. *Experimental Brain Research*, 40(2):237–240, 1980.

John A Hartigan and Manchek A Wong. Algorithm as 136: A k-means clustering algorithm. *Journal of the royal statistical society. series c (applied statistics)*, 28 (1):100–108, 1979.

Stefan Haufe, Vadim V Nikulin, Andreas Ziehe, Klaus-Robert Müller, and Guido Nolte. Combining sparsity and rotational invariance in eeg/meg source reconstruction. *NeuroImage*, 42(2):726–738, 2008.

Olaf Hauk. Keep it simple: a case for using classical minimum norm estimation in the analysis of eeg and meg data. *Neuroimage*, 21(4):1612–1621, 2004.

James V Haxby, Leslie G Ungerleider, Barry Horwitz, Jose Ma Maisog, Stanley I Rapoport, and Cheryl L Grady. Face encoding and recognition in the human brain. *Proceedings of the National Academy of Sciences*, 93(2):922–927, 1996.

H von Helmholtz. Ueber einige gesetze der vertheilung elektrischer ströme in körperlichen leitern, mit anwendung auf die thierisch-elektrischen versuche (schluss.). *Annalen der Physik*, 165(7):353–377, 1853.

Leonie Henschel, Sailesh Conjeti, Santiago Estrada, Kersten Diers, Bruce Fischl, and Martin Reuter. Fastsurfer-a fast and accurate deep learning based neuroimaging pipeline. *NeuroImage*, 219:117012, 2020.

Richard N Henson, Guillaume Flandin, Karl J Friston, and Jeremie Mattout. A parametric empirical bayesian framework for fmri-constrained meg/eeg source reconstruction. *Human brain mapping*, 31(10):1512–1531, 2010.

Richard N Henson, Daniel G Wakeman, Vladimir Litvak, and Karl J Friston. A parametric empirical bayesian framework for the eeg/meg inverse problem: generative models for multi-subject and multi-modal integration. *Frontiers in human neuroscience*, 5:76, 2011.

Ryan M Hill, Elena Boto, Molly Rea, Niall Holmes, James Leggett, Laurence A Coles, Manolis Papastavrou, Sarah K Everton, Benjamin AE Hunt, Dominic Sims, et al. Multi-channel whole-head opm-meg: Helmet design and a comparison with a conventional system. *NeuroImage*, 219:116995, 2020.

Arjan Hillebrand and Gareth R Barnes. The use of anatomical constraints with meg beamformers. *Neuroimage*, 20(4):2302–2313, 2003.

Arjan Hillebrand and Gareth R Barnes. Beamformer analysis of meg data. *International review of neurobiology*, 68:149–171, 2005.

Arjan Hillebrand and Gareth R Barnes. Practical constraints on estimation of source extent with meg beamformers. *Neuroimage*, 54(4):2732–2740, 2011.

Ana-Sofía Hincapié, Jan Kujala, Jérémie Mattout, Sebastien Daligault, Claude Delpuech, Domingo Mery, Diego Cosmelli, and Karim Jerbi. Meg connectivity and power detections with minimum norm estimates require different regularization parameters. *Computational intelligence and neuroscience*, 2016, 2016.

Joerg F Hipp, David J Hawellek, Maurizio Corbetta, Markus Siegel, and Andreas K Engel. Large-scale cortical correlation structure of spontaneous oscillatory activity. *Nature neuroscience*, 15(6):884–890, 2012.

Sepp Hochreiter and Jürgen Schmidhuber. Long short-term memory. *Neural computation*, 9(8):1735–1780, 1997.

Ming-Xiong Huang, Deborah L Harrington, Kim M Paulson, Michael P Weisend, and Roland R Lee. Temporal dynamics of ipsilateral and contralateral motor activity during voluntary finger movement. *Human brain mapping*, 23(1):26–39, 2004.

MX Huang, John C Mosher, and RM Leahy. A sensor-weighted overlapping-sphere head model and exhaustive head model comparison for meg. *Physics in Medicine & Biology*, 44(2):423, 1999.

Benjamin AE Hunt, Elizabeth B Liddle, Lauren E Gascoyne, Lorenzo Magazzini, Bethany C Routley, Krish D Singh, Peter G Morris, Matthew J Brookes, and Peter F Liddle. Attenuated post-movement beta rebound associated with schizotypal features in healthy people. *Schizophrenia bulletin*, 45(4):883–891, 2019.

RJ Ilmoniemi. The forward and inverse problems in the spherical model. *Biomagnetism: applications & theory*, 1985.

Ryouhei Ishii, Leonides Canuet, Ayako Ochi, Jing Xiang, Katsumi Imai, Derrick Chan, Masao Iwase, Masatoshi Takeda, O Carter Snead III, and Hiroshi Otsubo. Spatially filtered magnetoencephalography compared with electrocorticography to identify intrinsically epileptogenic focal cortical dysplasia. *Epilepsy research*, 81(2-3):228–232, 2008.

Ryouhei Ishii, Leonides Canuet, Tsutomu Ishihara, Yasunori Aoki, Shunichiro Ikeda, Masahiro Hata, Themistoklis Katsimichas, Atsuko Gunji, Hidetoshi Takahashi, Takayuki Nakahachi, et al. Frontal midline theta rhythm and gamma power changes during focused attention on mental calculation: an meg beamformer analysis. *Frontiers in human neuroscience*, 8:406, 2014.

Mark Jenkinson, Christian F Beckmann, Timothy EJ Behrens, Mark W Woolrich, and Stephen M Smith. Fsl. *Neuroimage*, 62(2):782–790, 2012.

Stephanie R Jones. When brain rhythms aren't 'rhythmic': implication for their mechanisms and meaning. *Current Opinion in Neurobiology*, 40:72–80, 2016.

Konstantinos Kamnitsas, Christian Ledig, Virginia FJ Newcombe, Joanna P Simpson, Andrew D Kane, David K Menon, Daniel Rueckert, and Ben Glocker. Efficient multi-scale 3d cnn with fully connected crf for accurate brain lesion segmentation. *Medical image analysis*, 36:61–78, 2017.

Meenakshi Khosla, Keith Jamison, Gia H Ngo, Amy Kuceyeski, and Mert R Sabuncu. Machine learning in resting-state fmri analysis. *Magnetic resonance imaging*, 64: 101–121, 2019.

Bjørg Elisabeth Kilavik, Manuel Zaepffel, Andrea Brovelli, William A MacKay, and Alexa Riehle. The ups and downs of beta oscillations in sensorimotor cortex. *Experimental neurology*, 245:15–26, 2013.

Yoon Kim. Convolutional neural networks for sentence classification. *arXiv preprint arXiv:1408.5882*, 2014.

Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.

Durk P Kingma, Tim Salimans, and Max Welling. Variational dropout and the local reparameterization trick. *Advances in neural information processing systems*, 28: 2575–2583, 2015.

Nadine Kloth, Christian Dobel, Stefan R Schweinberger, Pienie Zwitserlood, Jens Bölte, and Markus Junghöfer. Effects of personal familiarity on early neuromagnetic correlates of face perception. *European Journal of Neuroscience*, 24(11):3317–3321, 2006.

Amlan Kundu, Yang He, and Paramvir Bahl. Recognition of handwritten word: first and second order hidden markov model based approach. *Pattern recognition*, 22 (3):283–297, 1989.

Julian Kupiec. Robust part-of-speech tagging using a hidden markov model. *Computer speech & language*, 6(3):225–242, 1992.

M Lalancette, M Quraan, and D Cheyne. Evaluation of multiple-sphere head models for meg source localization. *Physics in Medicine & Biology*, 56(17):5621, 2011.

Joseph LeDoux. The amygdala. *Current biology*, 17(20):R868–R874, 2007.

Roberto Lent, Frederico AC Azevedo, Carlos H Andrade-Moraes, and Ana VO Pinto. How many neurons do you have? some dogmas of quantitative neuroscience under revision. *European Journal of Neuroscience*, 35(1):1–9, 2012.

Chin-Hsuan Lin, Tim M Tierney, Niall Holmes, Elena Boto, James Leggett, Sven Bestmann, Richard Bowtell, Matthew J Brookes, Gareth R Barnes, and R Chris Miall. Using optically pumped magnetometers to measure magnetoencephalographic signals in the human cerebellum. *The Journal of physiology*, 597(16):4309–4324, 2019.

Fa-Hsuan Lin, John W Belliveau, Anders M Dale, and Matti S Hämäläinen. Distributed current estimates using cortical orientation constraints. *Human brain mapping*, 27(1):1–13, 2006a.

Fa-Hsuan Lin, Thomas Witzel, Seppo P Ahlfors, Steven M Stufflebeam, John W Belliveau, and Matti S Hämäläinen. Assessing and improving the spatial accuracy in meg source localization by depth-weighted minimum-norm estimates. *Neuroimage*, 31(1):160–171, 2006b.

Vladimir Litvak, Alexandre Eusebio, Ashwani Jha, Robert Oostenveld, Gareth R Barnes, William D Penny, Ludvic Zrinzo, Marwan I Hariz, Patricia Limousin, Karl J Friston, et al. Optimized beamforming for simultaneous meg and intracranial local field potential recordings in deep brain stimulation patients. *Neuroimage*, 50(4): 1578–1588, 2010.

Jia Liu, Masanori Higuchi, Alec Marantz, and Nancy Kanwisher. The selectivity of the occipitotemporal m170 for faces. *Neuroreport*, 11(2):337–341, 2000.

Jia Liu, Alison Harris, and Nancy Kanwisher. Stages of processing in face perception: an meg study. *Nature neuroscience*, 5(9):910–916, 2002.

Quanying Liu, Marco Ganzetti, Nicole Wenderoth, and Dante Mantini. Detecting large-scale brain networks using eeg: impact of electrode density, head modeling and source localization. *Frontiers in neuroinformatics*, 12:4, 2018.

Yunzhe Liu, Raymond J Dolan, Zeb Kurth-Nelson, and Timothy EJ Behrens. Human replay spontaneously reorganizes experience. *Cell*, 178(3):640–652, 2019.

Lucrezia Liuzzi, Lauren E Gascoyne, Prejaas K Tewarie, Eleanor L Barratt, Elena Boto, and Matthew J Brookes. Optimising experimental design for meg resting state functional connectivity measurement. *Neuroimage*, 155:565–576, 2017.

José David López, William D Penny, Jairo J Espinosa, and Gareth R Barnes. A general bayesian treatment for meg source reconstruction incorporating lead field uncertainty. *NeuroImage*, 60(2):1194–1204, 2012.

Jose David Lopez, Vladimir Litvak, Jairo J Espinosa, K Friston, and Gareth R Barnes. Algorithmic procedures for bayesian meg/eeg source reconstruction in spm. *NeuroImage*, 84:476–487, 2014.

Henry T Luckhoo, Matthew J Brookes, and Mark W Woolrich. Multi-session statistics on beamformed meg data. *Neuroimage*, 95:330–335, 2014.

Eric Maris and Robert Oostenveld. Nonparametric statistical testing of eeg-and meg-data. *Journal of neuroscience methods*, 164(1):177–190, 2007.

Kanta Matsuura and Yoichi Okabe. Selective minimum-norm solution of the biomagnetic inverse problem. *IEEE Transactions on Biomedical Engineering*, 42 (6):608–615, 1995.

Hannah McCann, Giampaolo Pisano, and Leandro Beltrachini. Variation in reported human head tissue electrical conductivity values. *Brain topography*, 32(5):825–858, 2019.

Sofie S Meyer, James Bonaiuto, Mark Lim, Holly Rossiter, Sheena Waters, David Bradbury, Sven Bestmann, Matthew Brookes, Martina F Callaghan, Nikolaus Weiskopf, et al. Flexible head-casts for high spatial precision meg. *Journal of neuroscience methods*, 276:38–45, 2017a.

Sofie S Meyer, Holly Rossiter, Matthew J Brookes, Mark W Woolrich, Sven Bestmann, and Gareth R Barnes. Using generative models to make probabilistic statements about hippocampal engagement in meg. *Neuroimage*, 149:468–482, 2017b.

Tomáš Mikolov, Martin Karafiát, Lukáš Burget, Jan Černockỳ, and Sanjeev Khudanpur. Recurrent neural network based language model. In *Eleventh Annual Conference of the International Speech Communication Association*, 2010.

Akinori Nakamura, Takako Yamada, Atsuko Goto, Takashi Kato, Kengo Ito, Yuji Abe, Teruhiko Kachi, and Ryusuke Kakigi. Somatosensory homunculus as drawn by meg. *Neuroimage*, 7(4):377–386, 1998.

Walther Nernst. Die elektromotorische wirksamkeit der jonen. *Zeitschrift für physikalische Chemie*, 4(1):129–181, 1889.

Frank Neugebauer, Gabriel Möddel, Stefan Rampp, Martin Burger, and Carsten H Wolters. The effect of head model simplification on beamformer source localization. *Frontiers in neuroscience*, 11:625, 2017.

Guido Nolte. The magnetic lead field theorem in the quasi-static approximation and its use for magnetoencephalography forward calculation in realistic volume conductors. *Physics in Medicine & Biology*, 48(22):3637, 2003.

Magdalena Nowak, Emily Hinson, Freek Van Ede, Alek Pogosyan, Andrea Guerra, Andrew Quinn, Peter Brown, and Charlotte J Stagg. Driving human motor cortical oscillations leads to behaviorally relevant changes in local gabaa inhibition: a tacs-tms study. *Journal of Neuroscience*, 37(17):4481–4492, 2017.

Magdalena Nowak, Catharina Zich, and Charlotte J Stagg. Motor cortical gamma oscillations: what have we learnt and where are we headed? *Current Behavioral Neuroscience Reports*, 5(2):136–142, 2018.

George C O'Neill, Markus Bauer, Mark W Woolrich, Peter G Morris, Gareth R Barnes, and Matthew J Brookes. Dynamic recruitment of resting state sub-networks. *Neuroimage*, 115:85–95, 2015.

George C O'Neill, Prejaas K Tewarie, Giles L Colclough, Lauren E Gascoyne, Benjamin AE Hunt, Peter G Morris, Mark W Woolrich, and Matthew J Brookes. Measurement of dynamic task related functional networks using meg. *NeuroImage*, 146:667–678, 2017.

George C O'Neill, Ayan Sengupta, Michael Asghar, Eleanor L Barratt, Julien Besle, Denis Schluppeck, Susan T Francis, and Rosa M Sanchez Panchuelo. A probabilistic atlas of finger dominance in the primary somatosensory cortex. *NeuroImage*, 217: 116880, 2020.

George C O'Neill, Daniel N Barry, Tim M Tierney, Stephanie Mellor, Eleanor A Maguire, and Gareth R Barnes. Testing covariance models for meg source reconstruction of hippocampal activity. *bioRxiv*, 2021.

Robert Oostenveld, Pascal Fries, Eric Maris, and Jan-Mathijs Schoffelen. Fieldtrip: open source software for advanced analysis of meg, eeg, and invasive electrophysiological data. *Computational intelligence and neuroscience*, 2011, 2011.

Nobuyuki Otsu. A threshold selection method from gray-level histograms. *IEEE transactions on systems, man, and cybernetics*, 9(1):62–66, 1979.

Roberto D Pascual-Marqui, Christoph M Michel, and Dietrich Lehmann. Low resolution electromagnetic tomography: a new method for localizing electrical activity in the brain. *International Journal of psychophysiology*, 18(1):49–65, 1994.

Roger Penrose. A generalized inverse for matrices. 51(3):406–413, 1955.

Gavin Perry and Krish D Singh. Localizing evoked and induced responses to faces using magnetoencephalography. *European Journal of Neuroscience*, 39(9):1517–1527, 2014.

Usama Pervaiz, Diego Vidaurre, Chetan Gohil, Stephen M Smith, and Mark W Woolrich. Multi-dynamic modelling reveals strongly time-varying resting fmri correlations. *bioRxiv*, 2021.

Gert Pfurtscheller and FH Lopes Da Silva. Event-related eeg/meg synchronization and desynchronization: basic principles. *Clinical neurophysiology*, 110(11):1842–1857, 1999.

Robert Plonsey. The nature of sources of bioelectric and biomagnetic fields. *Biophysical Journal*, 39(3):309–312, 1982.

Mihai Popescu, Elena-Anda Popescu, Tszping Chan, Shannon D Blunt, and Jeffrey D Lewine. Spatio–temporal reconstruction of bilateral auditory steady-state responses using meg beamformers. *IEEE Transactions on biomedical engineering*, 55(3):1092–1102, 2008.

Andrew J Quinn, Diego Vidaurre, Romesh Abeysuriya, Robert Becker, Anna C Nobre, and Mark W Woolrich. Task-evoked dynamic network analysis through hidden markov modeling. *Frontiers in neuroscience*, 12:603, 2018.

Andrew J Quinn, Freek van Ede, Matthew J Brookes, Simone G Heideman, Magdalena Nowak, Zelekha A Seedat, Diego Vidaurre, Catharina Zich, Anna C Nobre, and Mark W Woolrich. Unpacking transient event dynamics in electrophysiological power spectra. *Brain topography*, 32(6):1020–1034, 2019.

Maher A Quraan, Sandra N Moses, Yuwen Hung, Travis Mills, and Margot J Taylor. Detection and localization of hippocampal activity using beamformers with meg: a detailed investigation using simulations and empirical data. Technical report, Wiley Online Library, 2011.

Lawrence Rabiner and Biinghwang Juang. An introduction to hidden markov models. *ieee assp magazine*, 3(1):4–16, 1986.

Muhammad Febrian Rachmadi, Maria del C Valdes-Hernandez, Maria Leonora Fatimah Agan, Carol Di Perri, Taku Komura, Alzheimer's Disease Neuroimaging Initiative, et al. Segmentation of white matter hyperintensities using convolutional neural networks with global spatial information in routine clinical brain mri with none or mild vascular pathology. *Computerized Medical Imaging and Graphics*, 66:28–43, 2018.

Stefan Rampp, Hermann Stefan, Xintong Wu, Martin Kaltenhäuser, Burkhard Maess, Friedhelm C Schmitt, Carsten H Wolters, Hajo Hamer, Burkhard S Kasper, Stefan Schwab, et al. Magnetoencephalography for epileptic focus localization in a series of 1000 cases. *Brain*, 2019.

Claire-Bénédicte Rivara, Chet C Sherwood, Constantin Bouras, and Patrick R Hof. Stereologic characterization and spatial distribution patterns of betz cells in the human primary motor cortex. *The Anatomical Record Part A: Discoveries in Molecular, Cellular, and Evolutionary Biology: An Official Publication of the American Association of Anatomists*, 270(2):137–151, 2003.

SE Robinson, SS Nagarajan, M Mantle, V Gibbons, and H Kirsch. Localization of interictal spikes using sam (g2) and dipole fit. *Neurology & clinical neurophysiology: NCN*, 2004:74, 2004.

SEVJ Robinson. Functional neuroimaging by synthetic aperture magnetometry (sam). *Recent advances in biomagnetism*, 1999.

Franck-Emmanuel Roux, Imène Djidjeli, and Jean-Baptiste Durand. Functional architecture of the somatosensory homunculus detected by electrostimulation. *The Journal of physiology*, 596(5):941–956, 2018.

Yannick Roy, Hubert Banville, Isabela Albuquerque, Alexandre Gramfort, Tiago H Falk, and Jocelyn Faubert. Deep learning-based electroencephalography analysis: a systematic review. *Journal of neural engineering*, 16(5):051001, 2019.

Berkman Sahiner, Heang-Ping Chan, Nicholas Petrick, Datong Wei, Mark A Helvie, Dorit D Adler, and Mitchell M Goodsitt. Classification of mass and normal breast tissue: a convolution neural network classifier with spatial domain and texture images. *IEEE transactions on Medical Imaging*, 15(5):598–610, 1996.

John G Samuelsson, Padmavathi Sundaram, Sheraz Khan, Martin I Sereno, and Matti S Hämäläinen. Detectability of cerebellar activity with magnetoencephalography and electroencephalography. *Human brain mapping*, 41(9):2357–2372, 2020.

Jukka Sarvas. Basic mathematical and electromagnetic concepts of the biomagnetic inverse problem. *Physics in Medicine & Biology*, 32(1):11, 1987.

Masa-aki Sato, Taku Yoshioka, Shigeki Kajihara, Keisuke Toyama, Naokazu Goda, Kenji Doya, and Mitsuo Kawato. Hierarchical bayesian estimation for meg inverse problem. *NeuroImage*, 23(3):806–826, 2004.

Peter H Schönemann. A generalized solution of the orthogonal procrustes problem. *Psychometrika*, 31(1):1–10, 1966.

Zelekha A Seedat, Andrew J Quinn, Diego Vidaurre, Lucrezia Liuzzi, Lauren E Gascoyne, Benjamin AE Hunt, George C O'neill, Daisie O Pakenham, Karen J Mullinger, Peter G Morris, et al. The role of transient spectral 'bursts' in functional connectivity: A magnetoencephalography study. *NeuroImage*, 209:116537, 2020.

Kensuke Sekihara and Srikantan S Nagarajan. *Electromagnetic brain imaging: a bayesian perspective*. Springer, 2015.

Kensuke Sekihara and Srikatan S Nagarajan. *Adaptive spatial filters for electromagnetic brain imaging*. Springer Science & Business Media, 2008.

Robert A Seymour, Nicholas Alexander, Stephanie J Mellor, George C O'Neill, Tim M Tierney, Gareth R Barnes, and Eleanor A Maguire. Using opms to measure neural activity in standing, mobile participants. *bioRxiv*, 2021.

Hyeyoung Shin, Robert Law, Shawn Tsutsui, Christopher I Moore, and Stephanie R Jones. The rate of transient beta frequency events predicts behavior across tasks and species. *Elife*, 6:e29086, 2017.

Alex Skates. *State-Switching Models of Human Brain Activity Using Recurrent Neural Networks (thesis)*. 2019.

Stephen M Smith, Peter T Fox, Karla L Miller, David C Glahn, P Mickle Fox, Clare E Mackay, Nicola Filippini, Kate E Watkins, Roberto Toro, Angela R Laird, et al. Correspondence of the brain's functional architecture during activation and rest. *Proceedings of the national academy of sciences*, 106(31):13040–13045, 2009.

Nelson Spruston. Pyramidal neurons: dendritic structure and synaptic integration. *Nature Reviews Neuroscience*, 9(3):206–221, 2008.

Jennifer R Stapleton-Kotloski, Robert J Kotloski, Jane A Boggs, Gautam Popli, Cormac A O'Donovan, Daniel E Couture, Cassandra Cornell, and Dwayne W Godwin. Localization of interictal epileptiform activity using magnetoencephalography with synthetic aperture magnetometry in patients with a vagus nerve stimulator. *Frontiers in neurology*, 5:244, 2014.

Matti Stenroos, Alexander Hunold, and Jens Haueisen. Comparison of three-shell and simplified volume conductor models in magnetoencephalography. *Neuroimage*, 94:337–348, 2014.

Michael C Stevens. The contributions of resting state and task-based functional connectivity studies to our understanding of adolescent brain network maturation. *Neuroscience & Biobehavioral Reviews*, 70:13–32, 2016.

Arjen Stolk, Ana Todorovic, Jan-Mathijs Schoffelen, and Robert Oostenveld. Online and offline tools for head movement compensation in meg. *Neuroimage*, 68:39–48, 2013.

M Streit, AA Ioannides, L Liu, W Wölwer, J Dammers, J Gross, W Gaebel, and H-W Müller-Gärtner. Neurophysiological correlates of the recognition of facial expressions of emotion as revealed by magnetoencephalography. *Cognitive brain research*, 7(4):481–491, 1999.

Vaanathi Sundaresan, Giovanna Zamboni, Nicola K Dinsdale, Peter M Rothwell, Ludovica Griffanti, and Mark Jenkinson. Comparison of domain adaptation techniques for white matter hyperintensity segmentation in brain mr images. *bioRxiv*, 2021.

Samu Taulu, Juha Simola, Jukka Nenonen, and Lauri Parkkonen. Novel noise reduction methods. *Magnetoencephalography: From Signals to Dynamic Cortical Networks*, pages 73–109, 2019.

Franca Tecchio, Patrizio Pasqualetti, Filippo Zappasodi, Mario Tombini, Domenico Lupoi, Fabrizio Vernieri, and Paolo Maria Rossini. Outcome prediction in acute monohemispheric stroke via magnetoencephalography. *Journal of neurology*, 254 (3):296–305, 2007.

CD Tesche and J Karhu. Theta oscillations index human hippocampal activation during a working memory task. *Proceedings of the National Academy of Sciences*, 97(2):919–924, 2000.

Tim M Tierney, Niall Holmes, Stephanie Mellor, José David López, Gillian Roberts, Ryan M Hill, Elena Boto, James Leggett, Vishal Shah, Matthew J Brookes, et al. Optically pumped magnetometers: From quantum origins to multi-channel magnetoencephalography. *NeuroImage*, 199:598–608, 2019.

Tim M Tierney, Andrew Levy, Daniel N Barry, Sofie S Meyer, Yoshihito Shigihara, Matt Everatt, Stephanie Mellor, Jose David Lopez, Sven Bestmann, Niall Holmes, et al. Mouth magnetoencephalography: A unique perspective on the human hippocampus. *NeuroImage*, 225:117443, 2021.

Andrey Nikolayevich Tikhonov. On the stability of inverse problems. In *Dokl. Akad. Nauk SSSR*, volume 39, pages 195–198, 1943.

Gerd Tinkhauser, Alek Pogosyan, Simon Little, Martijn Beudel, Damian M Herz, Huiling Tan, and Peter Brown. The modulatory effect of adaptive deep brain stimulation on beta bursts in parkinson's disease. *Brain*, 140(4):1053–1067, 2017.

Luzia Troebinger, José David Lopez, Antoine Lutti, David Bradbury, Sven Bestmann, and Gareth Barnes. High precision anatomy for meg. *Neuroimage*, 86:583–591, 2014.

Nelson J Trujillo-Barreto, David Araya, and Wael El-Deredy. The discrete logic of the brain-explicit modelling of brain state durations in eeg and meg. *bioRxiv*, page 635300, 2019.

Athina Tzovara, Sofie S Meyer, James J Bonaiuto, Aslan Abivardi, Raymond J Dolan, Gareth R Barnes, and Dominik R Bach. High-precision magnetoencephalography for reconstructing amygdalar and hippocampal oscillations during prediction of safety and threat. *Human brain mapping*, 40(14):4114–4129, 2019.

Freek van Ede, Andrew J Quinn, Mark W Woolrich, and Anna C Nobre. Neural oscillations: sustained rhythms or transient burst-events? *Trends in Neurosciences*, 41(7):415–417, 2018.

Barry D Van Veen and Kevin M Buckley. Beamforming: A versatile approach to spatial filtering. *IEEE assp magazine*, 5(2):4–24, 1988.

Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. Attention is all you need. In *Advances in neural information processing systems*, pages 5998–6008, 2017.

Diego Vidaurre, Andrew J Quinn, Adam P Baker, David Dupret, Alvaro Tejero-Cantero, and Mark W Woolrich. Spectrally resolved fast transient brain states in electrophysiological data. *Neuroimage*, 126:81–95, 2016.

Diego Vidaurre, Stephen M Smith, and Mark W Woolrich. Brain network dynamics are hierarchically organized in time. *Proceedings of the National Academy of Sciences*, 114(48):12827–12832, 2017.

Diego Vidaurre, Nicholas E Myers, Mark Stokes, Anna C Nobre, and Mark W Woolrich. Temporally unconstrained decoding reveals consistent but time-varying stages of stimulus processing. *Cerebral Cortex*, 29(2):863–874, 2019.

J Vrba and SE Robinson. Linearly constrained minimum variance beamformers, synthetic aperture magnetometry, and music in meg applications. In *Conference Record of the Thirty-Fourth Asilomar Conference on Signals, Systems and Computers (Cat. No. 00CH37154)*, volume 1, pages 313–317. IEEE, 2000.

J Vrba and SE Robinson. Squid sensor array configurations for magnetoencephalography applications. *Superconductor Science and Technology*, 15(9):R51, 2002.

Patrik Vuilleumier and Gilles Pourtois. Distributed and interactive brain mechanisms during emotion face perception: evidence from functional neuroimaging. *Neuropsychologia*, 45(1):174–194, 2007.

Daniel G Wakeman and Richard N Henson. A multi-subject, multi-modal human neuroimaging dataset. *Scientific data*, 2:150001, 2015.

Yi Wang, Hua Huang, Hao Yang, Jian Xu, Site Mo, Hongyu Lai, Ting Wu, and Junpeng Zhang. Influence of eeg references on n170 component in human facial recognition. *Frontiers in neuroscience*, 13:705, 2019.

David Wipf and Srikantan Nagarajan. A unified bayesian framework for meg/eeg source imaging. *NeuroImage*, 44(3):947–966, 2009.

Mark Woolrich, Laurence Hunt, Adrian Groves, and Gareth Barnes. Meg beamforming using bayesian pca for adaptive data covariance matrix regularization. *Neuroimage*, 57(4):1466–1479, 2011.

Mark W Woolrich, Adam Baker, Henry Luckhoo, Hamid Mohseni, Gareth Barnes, Matthew Brookes, and Iead Rezek. Dynamic state allocation for meg source reconstruction. *Neuroimage*, 77:77–92, 2013.

TT Yang, CC Gallen, BJ Schwartz, and FE Bloom. Noninvasive somatosensory homunculus mapping in humans by using a large-array biomagnetometer. *Proceedings of the National Academy of Sciences*, 90(7):3098–3102, 1993.

Tal Yarkoni, Russell A Poldrack, Thomas E Nichols, David C Van Essen, and Tor D Wager. Large-scale automated synthesis of human functional neuroimaging data. *Nature methods*, 8(8):665–670, 2011.

Heiga Zen and Haşim Sak. Unidirectional long short-term memory recurrent neural network with recurrent output layer for low-latency speech synthesis. In *Acoustics, Speech and Signal Processing (ICASSP), 2015 IEEE International Conference on*, pages 4470–4474. IEEE, 2015.

Rasmus Zetter, Joonas Iivanainen, Matti Stenroos, and Lauri Parkkonen. Requirements for coregistration accuracy in on-scalp meg. *Brain topography*, 31 (6):931–948, 2018.

Jinwei Zhang, Hang Zhang, Mert Sabuncu, Pascal Spincemaille, Thanh Nguyen, and Yi Wang. Bayesian learning of probabilistic dipole inversion for quantitative susceptibility mapping. In *Medical Imaging with Deep Learning*, pages 892–902. PMLR, 2020.

Yongyue Zhang, J Michael Brady, and Stephen Smith. Hidden markov random field model for segmentation of brain mr image. In *Medical Imaging 2000: Image Processing*, volume 3979, pages 1126–1137. International Society for Optics and Photonics, 2000.

Dongmei Zhi, Vince D Calhoun, Luxian Lv, Xiaohong Ma, Qing Ke, Zening Fu, Yuhui Du, Yongfeng Yang, Xiao Yang, Miao Pan, et al. Aberrant dynamic functional network connectivity and graph properties in major depressive disorder. *Frontiers in psychiatry*, 9:339, 2018.

Catharina Zich, Andrew J Quinn, Lydia C Mardell, Nick S Ward, and Sven Bestmann. Dissecting transient burst events. *Trends in Cognitive Sciences*, 24(10):784–788, 2020.