UNIVERSIDAD DE ANTIOQUIA

Brain-imaging based methodology for OPM sensor placement

Autor

Leonardo Duque Muñoz

Universidad de Antioquia

Facultad de Ingeniería, Doctorado en Ingeniería Electrónica

Medellín, Colombia

2019

Brain-imaging based methodology for

OPM sensor placement

Leonardo Duque Muñoz

Tesis Doctoral

como requisito para optar al título de:

Doctor en Ingeniería Electrónica

Asesores

Ph.D. Jose David Lopez Hincapie

PhD. Jesus Francisco Vargas Bonilla

Universidad de Antioquia

Facultad de Ingeniería.

Medellín, Colombia

2019.

Brain-imaging based methodology for OPM sensor placement

Leonardo Duque Muñoz

Universidad de Antioquia

Facultad de Ingeniería

Grupo de Investigación en Sistemas Embebidos e Inteligencia

Computacional SISTEMIC

October 2018

Brain-imaging based methodology for OPM sensor placement

Author:

Leonardo Duque Muñoz

Advisors:

José David López Hincapié

Jesús Francisco Vargas Bonilla

Universidad de Antioquia

Facultad de Ingeniería

Grupo de Investigación en Sistemas Embebidos

e Inteligencia Computacional SISTEMIC

November 2019

Declaration of Authorship

I, LEONARDO DUQUE MUÑOZ, declare that this thesis titled, 'Brain-imaging based methodology for OPM sensor placement' and the work presented in it are my own. I confirm that:

- This work was done wholly while in candidature for a research degree at this University.

- Where I have consulted the published work of others, this is always clearly attributed.

- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.

- I have acknowledged all main sources of help.

Leonardo Duque Muñoz

"The Road goes ever on and on
Down from the door where it began.
Now far ahead the Road has gone,
And I must follow, if I can,
Pursuing it with weary feet,
Until it joins some larger way
Where many paths and errands meet.
And whither then? I cannot say.

Still round the corner there may wait
A new road or a secret gate,
And though I oft have passed them by,
A day will come at last when I
Shall take the hidden paths that run
West of the Moon, East of the Sun."

J. R. R. Tolkien (Lord of the Rings books).

Abstract

Optically-pumped magnetometers (OPMs) have reached sensitivity levels that make them viable portable alternatives to traditional superconducting technology for magnetoencephalography. OPMs do not require cryogenic cooling, and can therefore be placed directly on the scalp surface. Unlike cryogenic systems based on a well characterised fixed arrays essentially linear in applied flux, or electroencephalography sensors that do not need to account for sensors orientation; OPM sensors are no longer rigidly arranged with a scanner system. Therefore, uncertainty in their locations and orientations with respect to the brain, and with respect to one another, must be accounted for. In this thesis dissertation, we propose a methodology to estimate the true sensor geometry of a disturbed array. We use parametric Bayesian inversion methods to perform neural source reconstruction and score among disturbed geometries with Free Energy as a cost function. This geometry disturbance is non-linear, causing local sub-optimal values on Free Energy that we tackle with a Metropolis search. Looking for a robust solution to this sensor placement problem, we develop a Multiple Kernel Learning (MKL) approach to extract the predominant complex dynamics hidden in the data. To do this, a weighted mixture of Gaussian kernels is used to highlight the data relationships, enhancing the data-driven covariance estimation and leading to a more reliable neural source reconstruction. When tested over disturbed OPM geometries, the MKL based solvers turned the Free Energy into a monotonic function, allowing the use of gradient descent optimisation. As a result, we estimate the true geometry of disturbed OPM arrays with a similar error than Metropolis search, but with 90% fewer iterations and allowing a larger search space. Our proposal suggests that a flexible and scalable design for sensor placement can be used to harness the potential of OPMs.

Acknowledgements

First, I would like to thank my advisors José David López Hincapié and Jesús Francisco Vargas Bonilla. Thank you both for the hours of mentoring, teaching and fruitful discussion. Thank you José for giving me the opportunity of improving my english skills (if by chance as reader you understand the written english in this thesis, you can be sure that it is because of José, if not, you can be sure that he did his best). Thank you for teaching me and for being so patient while I have been learning how to write, for finding me such an amazing place to do my research stay, the Wellcome Trust Centre for Human Neuroimaging, University College London.

This thesis would not have been possible without the help and support of Gareth R. Barnes, who I consider as an honorary advisor. He is one of the most kind people that I ever known, thank you for sharing your knowledge, your patience and for finding me financial support while I was in London, special thanks to Krish Singh for accepting me onto the partnership programme. I want to say thanks to many people from the WTCHN and Nottingham University, Tim Tierney, Sofie Meyer, James Bonaiuto, Sven Bestmann, Elena Boto, Niall Holmes, Gillian Roberts, James Legget, Richard Bowtell and Matthew Brookes, who have generously given me time, good advice, help, code and beers.

I started my PhD journey in the UDEA, I went to London, and my final year I did it while working in the Instituto Tecnológico Metropolitano. Thank you to my UDEA office mates who have kept me (somewhat) sane, Luis Germán, Andrés and Byron (I learned a lof of cyclism with you guys). Thank you to my current office mates and friends, Andrés E., Laura, Andrés F., Maria, Gloria and Hermes (thank you for understand that this year was crazy and be sure that my promise "after I finish my thesis" will be sustained, the moment is now).

I started four years ago, but to be honest, this started a while ago in Manizales, working with Germán Castellanos, and the GCPDS group. In there I learned how to collect and analise data, how to write a paper, many valuable things that I have used in this journey.

Many friends were there when I needed them, I have to thank Jorge Ivan, Paula, Angie, Juana, Ana Milena, Gustavo, Angela, Blanca Lucy, Gabriel, Oscar, Julian, Carlos. Thank you for sharing the weight. I have to specially give thanks to Juan David Martinez, he is not only my best friend, he also analised data, shared his knowledge and became part of this project. Juan and Paula made me godfather of Migue, who with his short life he gave me a final push to finish.

Thank you to my beloved family, my mother Francy, my father Cesar, my sister Luisa, my brothers Jacobo, Samuel and Juan Ignacio, and my grandma Iris. thank you for not loosing your faith in me, for always support me, for motivate me and share this journey with me. And thank you to my wife Ana Maria Duque for way too many things to list in one page.

Leonardo Duque Muñoz.

Contents

List of Figures

List of Tables

Abbreviations

BMA	**B**ayesian **M**odel **A**veraging
EBB	**E**mpirical **B**ayesian **B**eamformer
EEG	**E**lectroencephalography
GS	**G**reedy **S**earch
ISI	**I**nter **S**timulus **I**nterval
ITL	**I**nformation **T**eorethic **L**earning
KL	**K**ernel **L**earning
LOR	**L**oreta
MEG	**M**agnetoencephalography
MNI	**M**ontreal **N**eurological **I**nstitute
MKL	**M**ultiple **K**ernel **L**earning
MRI	**M**agnetic **R**esonance **I**maging
OPM	**O**ptically **P**umped **M**agnetometers
Rb	**R**ubidium
RDB	**R**eal **D**ata **B**ase
RKHS	**R**eproduced **K**ernel **H**ilbert **S**pace
ROI	**R**egion **O**f Interest
SDB	**S**imulated **D**ata Base
SEF	**S**omatosensory **E**voked cortical magnetic **F**ield
SERF	**S**pin **E**xchange **R**elaxation **F**ree regime
SK	**S**ingle **K**ernel
SNR	**S**ignal to **N**oise **R**atio
SPM	**S**tatistical **P**arametric **M**apping
SQUID	**S**uperconducting **Q**uantum Interference **D**evices
WMMK	**W**eighted combination of **M**ultiple Gaussian **K**ernels

Symbols

Notation

x	scalars
$\mathbf{x}, \mathbf{X}$	vectors, matrices
$\mathbf{X}_i$	i-th column of matrix $\mathbf{X}$
$\mathbf{X}_{ij}$	ij-th element of matrix $\mathbf{X}$
$\|\cdot\|_2$	L_2 norm or the Euclidean norm
$\|\cdot\|_F$	Frobenius norm
$E[\cdot]$	The expectation operator
$\mathcal{N}(\cdot)$	Gaussian distribution
$\mathcal{U}(\cdot)$	Uniform distribution
$\mathbf{X}^T$	Transpose of matrix $\mathbf{X}$
$\mathbf{X}^{-1}$	Inverse of matrix $\mathbf{X}$
$tr(\mathbf{X})$	Trace of matrix $\mathbf{X}$
$\mathbf{B}$	Magnetic field
$\mathbf{r}$	Position of the detector
$\mathbf{r_J}$	Position of a current source
$\mathbf{Y}$	EEG, MEG or OPM data
$\vec{y_i}$	Time-course of i-th MEG/EEG or OPM channel
$\mathbf{L}$	
$\mathbf{J}$	Current density
Ξ	Sensor noise

$\mathbf{Q}$	Prior source level covariance matrix
$\mathbf{Q}_\Xi$	Prior sensor noise covariance matrix
$\mathbf{Q_J}$	Posterior source covariance matrix
$\mathbf{Q_a}$	Model based sample covariance matrix of model a
$\mathbf{Q_Y}$	The data-based sample covariance matrix
N_c	Number of channels
N_t	Number of time samples
$\mathbf{D}_i$	i-th source covariance component matrix
$\vec{h}$	Vector of hyperparameters
$\mathbf{\Pi}$	Prior precision matrix of hyperparameters
Σ_h	Posterior covariance matrix of hyperparameters
F	Negative variational Free Energy
κ	Kernel function.
$\mathbf{K}$	Kernel covariance estimator.
σ	Kernel bandwidth
α	Weighting parameter
ξ	Prior location
θ	Positive stepsize parameter of gradient descent

For the two women that I love the most, my wife Ana María Duque Rojas, and my mother Francy del Socorro Muñoz Arboleda. Maybe they do not understand what is written in this document, but in the darkest times I found many answers in their eyes.

Chapter 1

Introduction

MEG/EEG Brain imaging techniques have become one of the most widely used methods for characterizing brain activity [1]. MEG/EEG Brain imaging consists of reconstructing the neural activity at cortical level from electromagnetic fields observed at scalp level [2]. Neural activity reconstruction is highly dependent on the propagation (forward) model used to relate cortical activity with scalp level recordings [3]. The spatial precision of the reconstruction increases with an appropriate model by for example including accurate information about the location and orientation of the sensors, or with a higher signal strength [4–8]. In this thesis, we propose to improve the quality of the information provided at both sensor and solver layers. We propose a methodology for selecting the best location and orientation of a new paradigm of MEG sensors called OPMs, by exploiting the non-stationary (and non-linear) nature of the recorded data.

Optically pumped magnetometers (OPMs) are new developed sensors to measure magnetic fields (MEG) at scalp level. The OPMs present an increase in the signal strength and a five fold improvement in the signal-to-noise-ratio (SNR) [6]. Nevertheless, being closer to the scalp, OPMs are prone to modelling errors [5]. As OPMs measure the same phenomena as traditional MEG, and MEG/EEG measurements are non-stationary and non-linear processes [9], it is expected that OPMs, being closer to the scalp and without a fixed sensor placement, would introduce new non-linearities to the measurement. Thus, the commonly used data covariance matrix may not be the best container of the MEG/EEG and OPM data, as it has several shortcomings such as being prone to be singular, having limited capability of modeling complicated relationships of the data (such as non-stationarities and non-linearities), and having a fixed form of representation [10]. In this thesis, we propose a new methodology for improving the forward models with a data driven optimization to estimate the location and orientation of the

OPMs sensor array, by using a multikernel approach for dealing with the non-linear and non-stationary nature of the neural-based electromagnetic data recordings.

1.1 Deling with OPM mis-localisation

OPMs provide a sensitive, flexible, and low-cost alternative to measure brain activity [6, 11]. OPMs do not require cryogenic cooling and can be placed flexibly on the scalp surface with a minimum separation of $\sim 4 - 7$ mm. OPMs potentially offer an increased sensitivity with respect to traditional MEG devices [12], and make the MEG system itself "wearable" with subjects able to move their head during the measurements [4]. In this way, OPMs have the potential to form the basis of a high signal SNR and flexible MEG instrumentation. However, this new flexibility introduces technical challenges such as the correct sensor placement and accurate modelling of the magnetic fields generated by the brain (i.e. accurate forward solutions) [6]. These issues must be overcome in order to create a practical, robust, and wearable system to be used for brain imaging.

The use of OPMs in wearable arrays brings uncertainty on both the absolute and relative sensor locations and orientations [5, 6, 13]. This contrasts with traditional MEG systems, where although there is uncertainty on the location of the head (which can be accounted for [14]; the relative channel locations and orientations are known with a high degree of accuracy [13]. For OPMs, these modelling errors have been minimized through the construction of MRI based subject-specific scanner-casts [5]. Such casts are 3D printed with predefined sensor slots and fit the subject specific head shape. The scanner-cast solution is useful for optimizing the data quality, as it removes a number of unknowns [15]; but it is not a practical solution, because it requires an investment per subject, and is both physically cumbersome and intimidating. Thus, in an ideal situation, one would like to use OPMs in flexible wearable arrays like those used for EEG electrodes, but with an accurate and robust methodology able to overcome the sensor placement issue (uncertainty about both the absolute and relative sensor locations and orientations).

Having an accurate model of the brain's anatomy and a correct spatial relationship between brain and sensors, the spatial specificity of the source reconstruction could be improved [13]. However, the reconstruction is highly dependent on the data's SNR. The dependence on accurate modelling is even more pronounced with OPMs because increments in SNR (due to the proximity of the sensor to the scalp) also entail increments in sensitivity to modelling errors [6]. In other words, if there is some topographical blurring in the data and large distance between sensors and sources, a small error in the propagation model makes little difference, but if the data have higher SNR and the

sensors are closer to the brain, small errors in the model lead to distorted estimates of the sources.

Some effort has been made to increase the spatial precision in brain imaging due to sensor placement uncertainty [14–18]. In [14], the authors proposed a methodology to recover the location of the cortical surface with poor prior knowledge of the head location within the MEG helmet space. The authors used a model comparison based methodology [19] combined with a Metropolis approach [20] to search over head positions. In [21], the authors developed a methodology to estimate the head position parameters dividing the MEG/EEG data into stationary sections and performing separate source inversions. These studies demonstrate that it is possible to estimate where the brain is within 4 mm based purely on MEG/EEG data, and to provide posterior distributions on source location that properly account for co-registration error.

Our aim in this work is to keep the spatial resolution of the scanner-cast, but in a flexible OPM array. Nevertheless, such work is not trivial as simulations in [6] showed that even small (5%) modelling errors could undermine the four-fold SNR increase with respect to traditional MEG promised by OPM systems. To do so, we make use of a cost function named Negative Variational Free Energy for model comparison [19] and two algorithms for performing the optimisation, the Metropolis search, and a gradient descent. The Free Energy rewards models that accurately fit the data, but penalizes models if they are overly complex (and therefore likely to overfit). However, this cost function relies on the data covariance matrix that, in practice is linearly estimated. This enables several shortcomings: the covariance matrix is prone to be singular, and has a limited capability to encode complicated relationships among data, with poor accuracy because of MEG/EEG non-stationarities and non-linearities [10]. To address these issues we propose a multi-kernel approach to account and exploit the hidden dynamics of the data.

1.2 Dealing with non-linearity of the data

To cope with the covariance matrix issues, in many machine learning tasks kernel methods have been proposed. These methods allow making use of object similarity metrics in a projected Reproduced Kernel Hilbert Space [22]. Thereby, kernel methods involve the use of positive definite matrices, which guarantee their non-singularity even if samples are scarce, as suitable object descriptors, that provide a solid framework for representing many types of data [23, 24]. Therefore, the kernel function is a flexible container for expressing knowledge about the problem as well as to capture the meaningful data relations[25].

Specifically, kernel-based methods are challenging due to the difficulty of choosing a proper Kernel for the problem at hand [25]. In this regard, among many available kernels, the Gaussian function is preferred since it aims to find a Reproducing Kernel Hilbert Space (RKHS) with universal approximation capability [26]. However, the proper estimation of the Gaussian Kernel bandwidth bounds the estimation of a RKHS, holding the main data relationships.

For tackling this issue, several authors have adjusted the kernel bandwidth based on the second-order statistics concepts, by exploiting the spatially varying data properties with local scales [27]. Other discussed approaches rely on Information Theoretic Learning (*ITL*) frameworks to quantify more broadly the data uncertainty [28]. However, either framework presents difficulties when the analyzed data comprise time-varying information because of the strong non-stationarity of MEG/EEG (and we expect OPM) data. Consequently, in order to relax the complex issue of fitting a single Kernel to the task at hand, several recent methods have focused on combining multiple kernels, which may correspond to different notions of similarity, namely, Gaussian Kernels with different bandwidth [29]. Thus, Multiple Kernel Learning (*MKL*) has become a powerful tool in absence of an appropriate data representation to highlight the leading data relationships.

In summary, to enhance the characterization of brain activity with OPMs and kernel functions, we propose a methodology to recover the sensor array location based on OPMs data. First, for traditional brain imaging techniques we evaluate Free Energy for different sensor placements over models. We then follow the same methodology but combining the traditional brain imaging approaches with a proposed multikernel approach that encodes different bandwidths to recover the predominant complex dynamics hidden in OPMs data.

The motivation of this thesis is to estimate the precise positions and orientations (and error bounds on these estimates) of an array of OPM sensors with their own recorded data. The methodology relies on a multi-kernel approach that exploit the non-stationary nature of the data to improve the forward model. If this is possible, then this procedure would reduce the dependence on 3D printed scanner-casts, suggesting that a more flexible and scalable design can be used to harness the potential of OPMs in a more practical manner. Moreover, it removes reliance on arbitrary scalp landmarks for co-registration, and provides an objective test of the quality of the data and forward models (i.e. whether they can be combined to recover the true OPM sensor locations).

1.3   Objectives

1.3.1   General Objective

To develop a methodology for characterizing brain activity based on feature extraction in source space, in order to implement it in optically pumped magnetometer systems.

1.3.2   Specific Objectives

- To analyze state-of-the-art techniques in EEG/MEG brain imaging, and propose a benchmark for testing the performance and interpretability of these techniques.

- To propose a robust methodology of feature extraction in source space, in order to bring interpretability while improving the performance of brain imaging systems.

- To select a validation benchmark for testing the proposed methodology and comparing it with state-of-the-art approaches.

1.4   Outline

The present research work is developed as follows: In Chapter 2 a background of the methods used in this thesis is presented. It includes a review of the MEG/EEG inverse problem solvers, and the use of Negative Variational Free Energy as a cost function to score among different models. Chapter 3 presents the methodology to find the position and orientation parameters of the sensors relative to the brain. Chapter 4 presents the theory behind the kernel functions applied to brain imaging, and a methodology to perform brain imaging with kernel priors. Chapter 5 presents the combination of both strategies to improve OPM sensor placement. Finally, general conclusions and main contributions of this research work are presented in Chapter 6.

In Chapter 3, we make use of real and simulated data from cryogenic multichannel recordings using a head-cast [15]; single channel OPM measurements using a scanner-cast [5]; and simultaneous multichannel measurements using the same scanner-cast. We then perturb our assumptions about the sensor positions and orientations obtaining a forward model for each hypothetical (or true) sensor configuration, and estimate the source distribution on the cortical surface by maximizing the model evidence over a range of sensor configurations. The model evidence is approximated by the negative variational Free Energy [30]. Each solution gives a single (maximal) Free Energy value for each possible sensor configuration. The sensor configuration that can provide the simplest explanation of the magnetic field produced by a current distribution on the individual's cortical surface will have the higher model evidence [19]. By using the scanner-cast (where relative sensor positions and orientations are known, and absolute positions and orientations are known to within ± 3 mm and ± 5 degrees respectively) with real measured data, we can directly test this method empirically.

In Chapter 4, to enhance the estimation of data covariance matrix, we propose a methodology to combine several kernels with different bandwidths for encoding the predominant complex dynamics hidden in MEG/EEG data. The methodology introduces a weighted mixture of Gaussian kernels using Multiple Kernel Learning, and allows mixing different notions of similarity to highlight the principal data relationships. Aiming to provide a proper source reconstruction accounting, we demonstrate to what extent a proper kernel combination enables increasing the performance of well-known distributed solutions to the inverse problem.

In Chapter 5, the strategy to estimate the OPM array geometry is combined with a solver that contains the multi-kernel implementation. When tested over disturbed OPM geometries, the MKL based solvers turned the Free Energy into a monotonic function, allowing the use of gradient descent optimisation. As a result, we estimate the true

geometry of disturbed OPM arrays with a similar error than Metropolis search, but with 90% fewer iterations and allowing a larger search space.

1.5 List of Publications

Papers in peer reviewed journals:

Paper 1: **L. Duque-Muñoz**, Tim M. Tierney, Sofie S Meyer, Elena Boto, Niall Holmes, Gillian Roberts, JF Vargas-Bonilla, Richard Bowtell, Matthew J. Brookes, JD López, Gareth R Barnes Data-driven model optimization to improve OPM co-registration (Published: Human Brain Mapping, 2019, volume 40, pages 4357-4369)

Paper 2: JD Martínez-Vargas, **L. Duque-Muñoz**, F. Vargas-Bonilla, JD Lopez, G. Castellanos-Domínguez Enhanced data covariance estimation using weighted combination of multiple Gaussian kernels for improved M/EEG source localization (Published: International Journal of Neural Systems, 2019, volume 29)

Paper 3: Tim M Tierney, Niall Holmes, Sofie S Meyer, Elena Boto, Gill Roberts, James Leggett, Sarah Buck, **L. Duque-Muñoz**, Vladimir Litvak, Sven Bestmann, Torsten Baldeweg, Richard Bowtell, Mathew J Brookes, Gareth R Barnes Cognitive neuroscience using wearable magnetometer arrays: Non-invasive assessment of language function (Published: Neuroimage, 2018, volume 181, pages 513-520).

Paper 4: Elena Boto, Niall Holmes, James Leggett, Gillian Roberts, Vishal Shah, Sofie S Meyer, **L. Duque-Muñoz**, Karen J Mullinger, Tim M Tierney, Sven Bestmann, Gareth R Barnes, Richard Bowtell Matthew J Brookes Moving brain imaging towards real-world applications using wearable MEG system (Published: Nature, 2018, volume 555, pages 657-661, doi:10.1038/nature26147)

Conference Papers:

Conference 1: Natural and Artificial Computation for Biomedicine and Neuroscience International Work-Conference on the Interplay Between Natural and Artificial Computation IWINAC 2017. Corunna, Spain. **L Duque-Muñoz**, JD Martínez-Vargas, G Castellanos-Domínguez, JF Vargas-Bonilla, JD Lopez- Non-linear Covariance Estimation for Reconstructing Neural Activity with MEG/EEG Data.

Conference 2: 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC 2016) **L. Duque-Muñoz**; F Vargas; JD López Simplified EEG inverse solution for BCI real-time implementation.

Conference 3: Congreso internacional de ciencias básicas e ingeniería, 2016. DM Marín-Castrillon, S Restrepo-Agudelo, HJ Areiza-Laverde, AE Castro-Ospina, **L Duque Muñoz**, Exploratory Analysis of Motor Imagery local database for BCI systems

Conference 4 Artificial Computation in Biology and Medicine International Work-Conference on the Interplay Between Natural and Artificial Computation IWINAC 2015, Elche, Spain. **L Duque-Muñoz**, RD Pinzón-Morales, G Castellanos-Domínguez EEG Rhythm Extraction Based on Relevance Analysis and Customized Wavelet Transform.

Conference 5: 6th Latin-American Conference on Networked and Electronic Media (LACNEM 2015) L Vega-Escobar, AE Castro-Ospina, **L Duque-Muñoz**, DWT-based feature extraction for motor imagery classification.

Conference 6: 20th Symposium on Signal Processing, Images and Computer Vision (STSIVA 2015). L. Vega-Escobar; AE Castro-Ospina, **L. Duque-Muñoz**, Feature extraction schemes for BCI systems.

Abstracts in conferences:

Conference 1: Workshop on Optically Pumped Magnetometers (BIOMAG 2018) **L-Duque-Muñoz**, JD López, JF Vargas-Bonilla, TM Tierney, R Bowtell, MJ Brookes, GR Barnes, Estimating the geometry of OPM sensor arrays relative to the human brain.

Conference 2: Workshop on Optically Pumped Magnetometers (WOPM-2017) **L-Duque-Muñoz**, JD López, JF Vargas-Bonilla, TM Tierney, R Bowtell, MJ Brookes, GR Barnes, Estimating the geometry of OPM sensor arrays relative to the human brain.

Conference 3: Workshop on Optically Pumped Magnetometers (WOPM-2017) Tim M. Tierney, Sofie Meyer, **L Duque-Muñoz**, JD ópez, James Osbourne, Vishal Shah, Richard Bowtell, Matthew J. Brookes, Gareth R. Barnes. Accuracy and Reliability of a multi-channel OPM MEG System for pre-surgical planning

Conference 4: MEG UK 2017 conference **L Duque-Muñoz**, JD Martínez-Vargas, G Castellanos-Domínguez, JF Vargas-Bonilla, JD López- Non-linear Covariance Estimation for Reconstructing Neural Activity with MEG/EEG Data.

Chapter 2

Background

In this chapter, an overview of magneto and electroencephalography (MEG/EEG) is presented, starting with the origins of MEG/EEG and two types of MEG sensors, namely superconducting quantum interference devices (SQUIDS) and optically pumped magnetometers (OPMs). Next, the theory and mathematical methods to characterize brain imaging and to perform model comparison are presented.

2.1 MEG and EEG

Electroencephalography (EEG) and magnetoencephalography (MEG) have been widely used to study brain dynamics, identifying and analyzing temporal activation patterns, e.g., neural rhythms [31, 32], event-related potentials (ERPs) [33], epileptic spikes [34], among others [35]. EEG and MEG are non-invasive techniques that measure electromagnetic brain activity with high temporal resolution (sampling below 1 ms. See [1, 36, 37] for a review in the area).

Modern SQUID sensors use a superconducting loop, which is usually made of niobium cooled to superconducting temperatures. The term cryogenic is therefore used to describe the MEG system's cooling which uses liquid helium to mantain a temperature of $\sim -270°$ C. SQUID sensors are most often arranged in a helmet-like structure or a dewar. The dewar contains hundreds of sensors distributed across its inner surface as close to the scalp as possible due to requirement for cryogenic cooling. Practically, MEG recording is silent, passive, non-invasive and gives exceptionally high time resolution (see [1] for a review in MEG as a tool for brain imaging).

2.1.1 Optically Pumped Magnetometers (OPMs)

Up to now, SQUIDS have remained as the only technology to measure brain magnetic fields. However, recent advances in atomic physics and miniaturization have let to introduction of a potential replacement technology. Optically Pumped Magnetometers (OPMs) are a new generation of MEG sensors which do not require cryogenic cooling. Instead of superconduction, OPMs rely on the spin polarization of alkali atoms [38].

Like SQUIDs, OPMs development began nearly 50 years ago [39], but initially with a dramatically larger size and power consumption, as well as poorer sensitivity. Particularly because of their size, these magnetometers could not be used for multichannel recordings, making them less attractive for MEG experiments. However, over the past decade, these problems have been solved and now the technology represent a candidate for replacing SQUIDs. The primary reason is that OPMs have been miniaturized and operate without cryogenic cooling, meaning that they can be placed directly on the scalp. Additionally, they provide equal sensitivity to magnetic fields than SQUIDs [12]. The OPMs are commercially available http://quspin.com.

These devices rely on the Spin Exchange Relaxation Free (SERF) regime [40], which can occur in very low magnetic fields, when the Lamour frequency is considerably lower than spin-spin interaction. These sensors have three crucial components: a laser (795 nm wavelength), a Rb^{87} vapour cell, and a photodiode. The laser pumps the Rb^{87} vapour into steady state with large atomic polarization. In this state the vapour becomes transparent to the laser light allowing the light to pass through the vapour with limited loss of energy and therefore a maximum detection at the photodiode. However, the polarization is highly sensitive to the ambient magnetic field. As the ambient magnetic field changes (due to brain's activity) the polarization and transparency of the vapour is reduced. This means that more energy is absorbed by the gas (from the laser light) when the ambient magnetic field changes. As such, the magnetic field modulates how much light is detected at the photodiode. Here, the OPMs are configured to produce a sinusoidal oscillating ($\sim$100 Hz) magnetic field to modulate the signal along two axes. This allows measuring both radial and tangential (to the head) components of the magnetic field. The interested reader is directed to other general overviews of the physical principles of OPMs [38, 39, 41–43].

The QuSpin OPM sensors used here have a noise level comparable to SQUIDs (~ 15 fT$/\sqrt{Hz}$ above 10 Hz), a bandwidth up to 130 Hz (1st order cut-off), an operational dynamic range of ±1.5 nT), a size of $14 \times 21 \times 80$ mm^3, and can be placed such that the sensitive volume is 6.5 mm from the scalp surface. We modelled the sensitive volume of gas as a single point measurement of the field normal to the sensor base.

2.1.2 OPM vs SQUIDs

Although OPMs have a noise level comparable to SQUIDs, being closer to the head allows improving the magnitude of the measured MEG signal. This improvement afforded by moving the detector closer to the head is independent on the location in the head of the source generating the signal.

Considering a dipole oriented normal to the radial field direction (specifically, the z-direction), the magnetic field $\mathbf{B}$, at $\mathbf{r}$ (position of the detector, SQUIDS or OPMs) is given by:

$$\mathbf{B}_t(\mathbf{r}) = -\frac{\mu_0 \left|\mathbf{J}\right| \left|\mathbf{r_J}\right| sin(\phi)}{4\pi \left(\left|\mathbf{r}\right|^2 + \left|\mathbf{r_J}\right|^2 - 2\left|\mathbf{r}\right|\left|\mathbf{r_J}\right| cos(\phi)\right)^{3/2}} \tag{2.1}$$

Eq. (2.1) was derived from the Geselowits formula (see [5, 44] for its demonstration), where $\mathbf{r_J}$ is the position of a current source $\mathbf{J}$ inside a volume conductor. The expression in Eq. (2.1) is asymmetric with respect to the angle ϕ (between $\mathbf{r}$ and $\mathbf{r_J}$), and the angle at which the maximum field is found varies as a function of both $\left|\mathbf{r}\right|$ and $\left|\mathbf{r_J}\right|$. The optimal angle at which sensor should be placed is found by differentiating Eq. (2.1), so that:

$$cos(\phi) = \frac{-(\left|\mathbf{r}\right|^2 + \left|\mathbf{r_J}\right|^2) + \sqrt{\left|\mathbf{r}\right|^4 + \left|\mathbf{r_J}\right|^4 + 14\left|\mathbf{r}\right|^2 \left|\mathbf{r_J}\right|^2}}{2\left|\mathbf{r}\right|\left|\mathbf{r_J}\right|^2} \tag{2.2}$$

Fig. 2.1 presents the Magnitude of the measured radial magnetic field as a function of source depth of SQUID (red) and OPM (blue). Note that the OPM offers an improvement in source magnitude for sources anywhere in the brain.
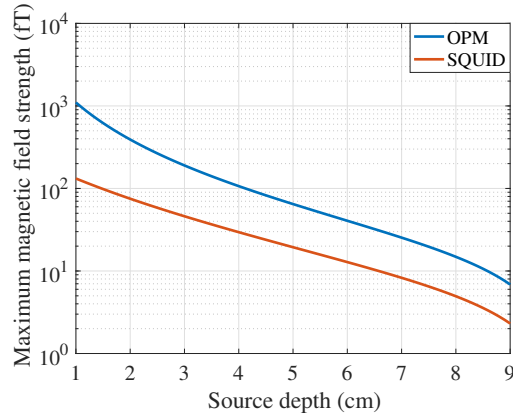
FIGURE 2.1: Magnitude of the measured radial magnetic field as a function of source depth for SQUID and OPM. Note that the improvement factor in magnetic field strength in superficial sources could be as high as eight times the magnitude, while for depth sources (6 cm) the improvement could be of three times in magnitude.

2.1.3 EEG Vs MEG

Table 2.1 presents a comparative analysis between EEG and MEG with both type of sensors SQUIDs and OPMs. It shows that the OPMs have a better spatial resolution, allows movement of the subjects, it is wearable, and easier to study young populations. However, as it measures magnetic fields, there is a need of a special shielded room which increases its cost. However, in [4] an array of coils that suppresses the remnant field in the shielded room was presented. Theoretically, in a future the sensors could be used without the need of the shielded room as the OPMs could have an active noise cancellation combined with such coils. As the OPMs have the best of EEG and MEG, the advantages of OPMs seems promising for the future of brain imaging.

TABLE 2.1: EEG and MEG modalities comparison.

EEG	SQUIDs	OPM
Signal distorted by skull-scalp	Signal unaffected by skull-scalp	Signal unaffected by skull-scalp
Spatial resolution $\sim 7 - 16$ cm	Spatial resolution $\sim 3 - 11$ mm	Spatial resolution $\sim 1 - 7$ mm
Temporal Resolution 1 ms	Temporal Resolution 1 ms	Temporal Resolution 10 ms
Allow subjects to move	Subjects must remain still	Allows subjects to move
Flexible location of sensors attached to the head	Sensors fixed to a helmet $\sim 3 - 6$ cm to the head	Sensors in fixed arrays close to the head (~ 6 mm)
Wearable	-	Wearable
Can be done anywhere	Requires special magnetic shielded room	Requires special magnetic shielded room
Studies of young population feasible	Fixed helmet designed for adults	Studies of young young population feasible

2.2 MEG/EEG Inverse Problem

An externally measured electromagnetic field can be used to estimate neural activity in terms of neural source distribution and time course. This requires two steps: construction of a forward model to describe the predicted scalp distributions for each given source of neural activity with a specific orientation, magnitude and location; and subsequent inversion of this forward model to estimate the most likely spatial configuration of sources giving rise to the measured signal. All the inversion methods used here are parametric Bayesian inverse methods. The definitions and implications thereof will be described in this section. All source reconstruction methods rely on carrying out the following steps: preprocessing, including filtering and removal of possible artifacts present in the data; source space modelling; data co-registration; forward model computation; and source space reconstruction. This thesis is focused on the latter four. The methods for source reconstruction vary with respect to assumptions about source space covariance.

The electromagnetic field measured with MEG/EEG may be represented by the general linear model [45].

$$\mathbf{Y} = \mathbf{L}\mathbf{J} + \varXi \tag{2.3}$$

Where $\mathbf{Y} \in \mathbb{R}^{N_c \times N_t}$ are the measured MEG/EEG data with N_c channels and N_t time samples affected by zero mean Gaussian noise distribution $\varXi = \mathcal{N}(0, \mathbf{Q}_\varXi)$, $\varXi \in \mathbb{R}^{(N_c \times N_d)}$, and noise covariance $\mathbf{Q}_\varXi \in \mathbb{R}^{N_c \times N_t}$. $\mathbf{J} \in \mathbb{R}^{N_d \times N_t}$ is the current flow due to N_d current dipoles distributed across the cortical surface, with prior Gaussian assumptions $\mathbf{J} = \mathcal{N}(0, \mathbf{Q})$. $\mathbf{Q} \in \mathbb{R}^{N_d \times N_d}$ is the source level covariance matrix. The gain matrix $\mathbf{L} \in \mathbb{R}^{(N_c \times N_d)}$ (commonly known as lead-field matrix) is defined as the propagation model of an MEG signal that is produced by a source of unitary strength, and is completely determined by the sensor configuration and volume conductor physics, as described in terms of the forward model.

In the linear model Eq. (2.3), $\mathbf{L}$ is non invertible since the dipoles outnumber the sensors ($N_d \gg N_c$). Then, the reconstruction of the sources of neural activity $\widehat{\mathbf{J}}$ cannot be performed directly. The basis of the Bayesian implementation is that the recorded activity over sensors $\mathbf{Y}$ is used to estimate the distribution of sources in the brain. The constraints take form of prior probabilities of source activity $p(\mathbf{J})$ and these priors are informed by anatomy and neurophysiology (and combined with the physical properties of volume conduction). The priors are then used to estimate the posterior probability of the source activity due to data $P(\mathbf{J}|\mathbf{Y})$, allowing to estimate the posterior source

activity distribution using the Bayes' theorem:

$$p(\mathbf{J}|\mathbf{Y}) = \frac{p(\mathbf{Y}|\mathbf{J})p(\mathbf{J})}{p(\mathbf{Y})} \tag{2.4}$$

Where evidence for the recorded data $p(\mathbf{Y})$ is considered to be fixed, given a constant dataset. The estimated magnitude of $\mathbf{J}$, $\widehat{\mathbf{J}}$ can be computed by taking the expectation of the posterior $\widehat{\mathbf{J}} = E[p(\mathbf{J}|\mathbf{Y})]$. Typically, the noise associated to MEG/EEG measurements is considered white Gaussian: $p(\varXi) = \mathcal{N}(\varXi; 0, \mathbf{Q}_\varXi)$. Under this assumption the likelihood is defined Gaussian as well.

In this work, the framework presented in [14] is adapted to explore among perturbed arrays. Briefly, with each perturbed array a new lead-field matrix $\mathbf{L}_a$ (a-th forward model) is computed using the single shell forward model [46] for MEG data, and a BEM model for EEG. For uninformative priors, the Maximum-likelihood solution to the inverse problem reduces to:

$$\widehat{\mathbf{J}} = \mathbf{Q}\mathbf{L}_a^T(\mathbf{Q}_\varXi + \mathbf{L}_a\mathbf{Q}\mathbf{L}_a^T)^{-1}\mathbf{Y} \tag{2.5}$$

2.2.1 Prior Information

Intending to enhance estimations of $\mathbf{Q}$, a large variety of priors have been proposed within the MEG/EEG inverse problem, which are mostly divided into two approaches [47]: model-based (or anatomically-based) and data-driven. The former strategy patterns the synchronized activation of brain pyramidal cells (smoothness condition) as it is the case of Low-Resolution brain electromagnetic tomography (LOR) that includes a Laplacian operator $\nabla \in \mathbb{R}^{N_d \times N_d}$ (or smoother) to represent the neuron groups with synchronized activation:

$$\mathbf{Q} = (\mathbf{L}^T\nabla^T\nabla\mathbf{L})^{-1} \tag{2.6}$$

In data-driven approaches (namely, Beamforming [48]), the estimates of $\mathbf{Q}$ are performed straightforwardly by projecting the available MEG/EEG recordings into the source space through the lead-field matrix. This is the case of Empirically Bayesian Beamformers (EBB), which generates a single covariance diagonal matrix with elements computed by projecting the measured data covariance into another source space, as follows [49]:

$$\mathbf{Q} = \mathrm{diag}((\boldsymbol{l}_i^T\mathbf{Q}_\mathbf{Y}^{-1}\boldsymbol{l}_i)^{-1}/\delta_i); \ \forall i = 1, \ldots, N_d \tag{2.7}$$

where $l_i \in \mathbb{R}^{N_c \times 1}$ is i-th column of $\mathbf{L}$, matrix $\mathbf{Q}_Y \in \mathbb{R}^{N_c \times N_c}$ is the MEG/EEG linear data covariance computed as $\mathbf{Q_Y} = \mathbf{YY}^T/N_t$, and the normalisation parameters $\delta_i \in \mathbb{R}^+$ are fixed as $\delta_i = 1/l_i^T l_i$.

However, both estimation approaches (LOR and EBB) may face limited ability to generalize due to their linear representation premise, which cannot be expanded as to model nonlinear feature relationships [50], or to include additional prior information. But being formed by independent priors, $\mathbf{Q}$ may be relaxed as a linear combination of N_p components [51]: $\mathbf{Q} = \sum_{i \in N_p} h_i \mathbf{D}_i$, where the weight collection of $\{h_i \in \mathbb{R}^+\}$ determines the contribution of a given dictionary $\mathcal{D} = \{\mathbf{D}_1, \ldots, \mathbf{D}_{N_p}\}$. Each matrix $\mathbf{D}_i \in \mathbb{R}^{N_d \times N_d}$ can embody priors extracted from previous approaches and/or even other modalities such as functional magnetic resonance images [52].

In this work we use Multiple Sparse Priors (MSP [19]) implementation, for which the hyperparameters h_i are estimated by an Expectation-Maximization optimization using the negative variational Free Energy as the cost function. Intending to reach a trade-off between accuracy, $f_{AC} \in \mathbb{R}$, and complexity, $f_{CX} \in \mathbb{R}$. Thus, the MSP inverse solution employs the negative variational Free Energy as cost function [30], defined as $F = f_{AC} - f_{CX}$.

2.3 Negative variational Free Energy

Each inversion returns a Negative Variational Free Energy value (F), which approximates the model evidence $p(\mathbf{Y}|a)$ (where a is the model) [30, 51]. In this thesis, we use the Free Energy to score competing source reconstructions based on different sensor locations and orientations (different geometries modelled through different $\mathbf{L}_a$ models). That is, reconstructions of the same data but with different sensor configurations, each providing an associated Free Energy that can be compared across geometries [53]. For a model $\mathbf{L}_a$ associated with a given sensor location and orientation, Free Energy F_a can be expressed as a trade-off between accuracy and complexity:

$$F_a = f_{AC}(a) - f_{CX}(a) \tag{2.8}$$

The accuracy term is expressed as:

$$f_{AC}(a) = \frac{N_c}{2} trace(\mathbf{Q_Y}\mathbf{Q}_a^{-1}) - \frac{N_c}{2} log\,|\mathbf{Q}_a| - \frac{N_c N_t}{2} log(2\pi) \tag{2.9}$$

Where $\mathbf{Q_Y} = (\frac{1}{N_c}\mathbf{YY}^T)$ the data-based sample covariance matrix, N_t is the number of samples, and $|\cdot|$ is the matrix determinant operator. When searching for the optimal

geometry, the MEG/EEG data do not change; so, the accuracy of the model a mainly depends on the model-based sample covariance matrix computed as $\mathbf{Q}_a = \mathbf{Q}_\Xi + \mathbf{L}_a \mathbf{Q} \mathbf{L}_a^T$.

The complexity term depends on the hyperparameters h, which provide a trade-off between sensor noise $\mathbf{Q}_\Xi = h_1 \mathbf{I}(N_c)$, and the Beamforming prior $\mathbf{Q}_a = h_2 \mathbf{\Gamma}$ (with $\mathbf{\Gamma}$ being the beamformer prior):

$$f_{CX}(a) = \frac{1}{2}(\widehat{\boldsymbol{\lambda}}_a - \boldsymbol{\nu})^T \mathbf{\Pi} (\widehat{\boldsymbol{\lambda}}_a - \boldsymbol{\nu}) + \frac{1}{2} \log |\mathbf{\Sigma}_{\boldsymbol{\lambda}_a} \mathbf{\Pi}| \tag{2.10}$$

The prior and posterior distributions of the hyperparameters are considered Gaussian: $q(\boldsymbol{\lambda}_a) = \mathcal{N}(\boldsymbol{\lambda}; \boldsymbol{\nu}, \mathbf{\Pi}^{-1})$ and $p(\boldsymbol{\lambda}_a) = \mathcal{N}(\hat{\boldsymbol{\lambda}}_a, \mathbf{\Sigma}_{\boldsymbol{\lambda}_a})$, respectively (where $\hat{\boldsymbol{\lambda}}_a$ and $\mathbf{\Sigma}_{\lambda_a}$ are the posterior mean and covariance of the hyperparameters for model a). We used the standard SPM https://www.fil.ion.ucl.ac.uk/spm/software/spm12/ implementation of this algorithm with non-informative mean and precision ($\boldsymbol{\nu}$ and $\mathbf{\Pi}$) [54].

As the Free Energy approximates the model evidence for inversions obtained with different models (different geometries) but same data; it can be used to compare the likelihood of these models. In this model comparison framework, the $F(a)$ values obtained for different geometries can be compared with respect to how well they describe the data by simply substracting one F value from the other. It is thereby possible to quantify the difference in (approximated) model evidence. Because of the Bayesian context, the F value difference is equivalent to calculating a Bayes factor. In line with Bayesian convention, a positive difference means that the first model in the equation $\frac{1}{e^{-\Delta F}}$ is more likely than the second. A significance threshold is defined at three, due to the log distribution of the Bayes factor of F difference, one model is 20 times more likely than the other. Critically, model comparison is only valid when the data is the same and can only be used to infer relative fitness of different models. The application of this model comparison aproach to MEG data analysis has been succesfully demostrated elsewhere [14, 16, 54–57].

Chapter 3

Data-driven model optimisation to improve OPM co-registration

3.1 Introduction

Optically-pumped magnetometers (OPMs) have reached sensitivity levels that make them viable portable alternatives to traditional superconducting technology for magnetoencephalography (MEG) [4, 58]. The devices have a footprint of only 2.47 cm^2, do not require Helium cooling, and can therefore be placed directly on the scalp surface, increasing sensitivity to the brains magnetic fields. However, this flexibility in sensor placement introduces a new source of spatial uncertainty which limits the accuracy of neuronal current flow estimates. Specifically, since sensors are no longer rigidly arranged inside a scanner system, their locations and orientations with respect to the brain, and respect to one another, must be accounted for in the forward model. In this chapter, we perturb the sensor geometry (via simulation) and with analytic model comparison methods estimate the true sensor geometry. The resulting curves allows us to compare different MEG systems. We test this technique using simulated and real data from SQUID and OPM recordings using head-casts and scanner-casts. We demonstrate how this framework can also be used to optimise single sensor and sensor array models. Finally, we show that given knowledge of underlying brain anatomy, it is possible to estimate the true sensor geometry from the OPM data themselves using a model comparison framework. This implies that the requirement for accurate knowledge of the sensor positions and orientations a priori may be relaxed. As this procedure uses the cortical manifold as spatial support there is no co-registration procedure or reliance on arbitrary scalp landmarks.

To carry out our analysis, we make use of real and simulated data from cryogenic multichannel recordings using a head-cast [15]; single channel OPM measurements using a scanner-cast[5]; and simultaneous multichannel measurements using the same scanner-cast. We then perturb our assumptions about the sensor positions and orientations obtaining a forward model for each hypothetical (or true) sensor configuration, and estimate the source distribution on the cortical surface by maximizing the model evidence over a range of sensor configurations.

3.2   Methodology

The methodology used to recover the geometry of the array is presented in Fig. 3.1. The methodology comprises the following steps. *i)* OPM data collection. *ii)* Variation in the array geometry (the algorithm do not know the true position). *iii)* Forward model computation, the single shell model was used to compute the propagation model. *iv)* Estimation of the most likely cortical current distribution with the perturbed model. *v)* Metropolis search procedure with Free Energy as a cost function (repetition of steps *iv)* and *v)* until convergence. *vii)* Bayesian model averaging that pool estimates from across a range of optimisation steps and weight them by their model evidence (see Apendix A for the Metropolis search and BMA algorithms).
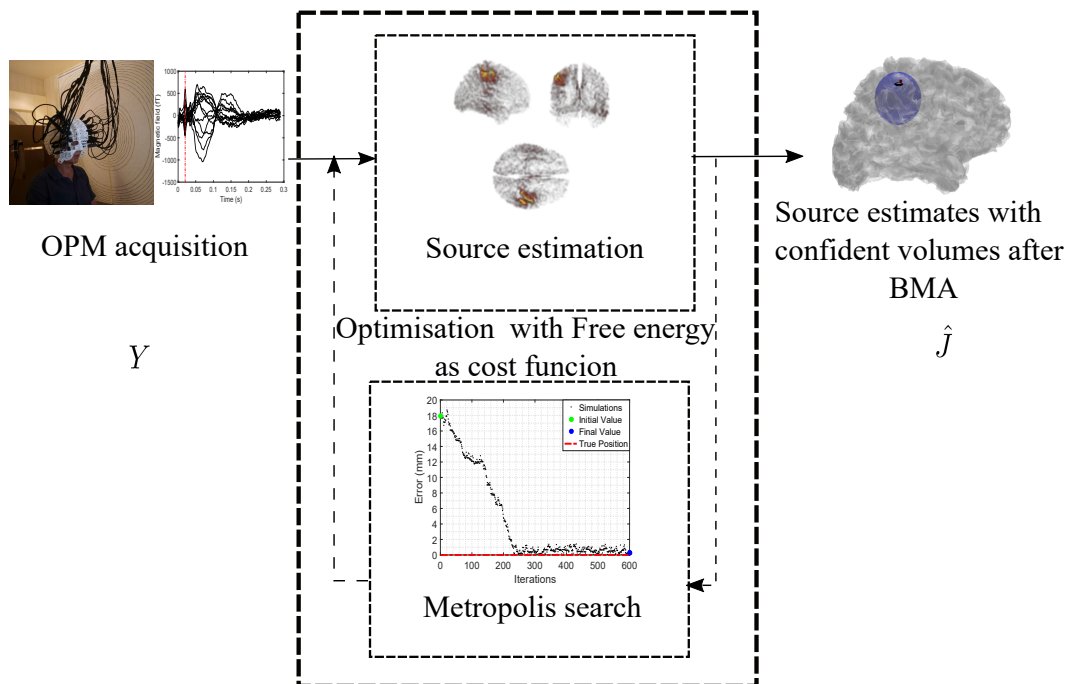
FIGURE 3.1: Benchmark used to recover the true geometry of the OPM array. The location and orientation parameters that minimises the error in the Free Energy cost function are the most likely ones.

3.3 Experimental Setup

3.3.1 Data collection

We used empirical data from a somatosensory evoked response paradigm which involved electrical stimulation of the subjects left median nerve. There are three data sets collected from the same subject and using the same paradigm used in this paper: i) data collected with a SQUID system, ii) using multiple repeats of the same experiment with a single OPM channel at different locations, and iii) using an array of 13 OPM channels operating simultaneously Fig. 3.2(b). The data and recording parameters for i) and ii) have already been described in [5].

Briefly, we performed a left median nerve electrical stimulation by applying a series of 500 μs duration current pulses to two gold electrodes placed on the subject's left wrist. The current was applied using a Digitimer DS7A constant current stimulator, and the amplitude was increased until a visible movement of the thumb was observed upon stimulation. Each experimental run comprised 80 pulses delivered with an inter-stimulus-interval (ISI) of 3.0099 s; the ISI was selected not to be a multiple of the 20 ms period of the mains noise frequency (50 Hz). Data were recorded during a prestimulus baseline of 0.5 s and post stimulus period of 2 s. A single run lasted four minutes.

All recordings were carried out inside a magnetic shielded room comprising two layers of mu-metal and one of aluminium. For the OPM recordings, the subject was positioned on the bed with his head in the scanner-cast and the scanner-cast was rigidly fixed in position Fig. 3.2(b). The single OPM measurements made with a sequential sampling of scanner-cast slots using a single OPM channel are explained in [5]. Here we additionally ran the experiment while 13 sensors located in the same slots of the scanner-cast were operating simultaneously. The scanner-cast presented in Fig. 3.2(a) was used for both sets of OPM recordings. OPM data were acquired simultaneously with SQUID data (from a 275 channel CTF instrument), and the magnetometer reference channels within SQUID system (remote from the subject) and the time-derivatives of these channels were used as an environmental noise reference set and regressed out of the OPM data on a trial by trial basis (as described in [5]. SQUID recordings were performed using the 275 channel system in 3rd gradient configuration (i.e. with factory-set linear weighting from the noise reference array).

3.3.2 Scanner-casts

As a basis for both the simulated and empirical experiments we used the array geometry as defined in [5]. Briefly, this relies on 3D printing to construct an individualized helmet containing a sensor array positioned over the subjects sensory motor cortex (Fig. 3.2(a), for more details, see [5]). As the scanner-cast was built directly from the subjects MRI, the location and orientation of the cast with respect to the brain anatomy was known to within ± 3 mm and ± 5 degrees (conservative estimates based on how far the cast could be manipulated whilst on the subject).
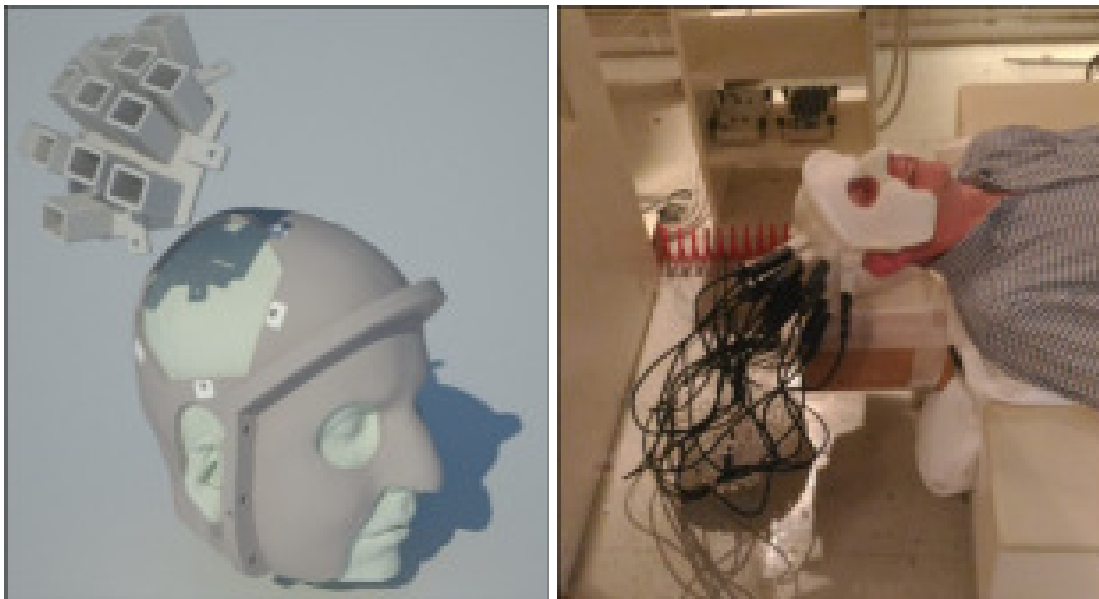
(a) Scanner-cast   (b) OPM recording set-up

FIGURE 3.2: Head Cast and OPM recordings.3.2(a) CAD model of the scanner-cast. The cast is based on an individual MRI scan and designed to house the OPM sensors around the outer scalp surface.3.2(b) Subject wearing the scanner-cast with the multichannel OPM array. The scanner-cast is fixed rigidly within the shielded room eliminating any sensor motion.

3.3.3 Simulated data

The SPM12 https://www.fil.ion.ucl.ac.uk/spm/software/spm12/ software package is used to simulate single trial MEG datasets based on both the OPM and SQUID sensor geometries used empirically. For simplicity and comparability, all the simulated trials had the same duration (1 s) and number of sensor channels ($N_c = 13$). We simulated a single 10Hz sinusoidal source located in the somatosensory cortex (at 46, -25, 60 mm in MNI space) with a dipole moment of 10 nAm. We added the same Gaussian white background noise of standard deviation 100 fT RMS to both SQUID and OPM

sensors, we kept background noise levels equal to make interpretation of both systems easier.

3.3.4 Variation in the array geometry

To assess whether we could derive the correct sensor geometry based on the OPM data, we perturbed the sensor array in two ways. Firstly, we randomly perturbed the orientation of each sensor independently within the OPM array. For each sensor, the axis of the perturbation (roll, pitch or yaw in x, y or z) was selected randomly and these perturbations were moved in 2.5 degree steps between -20 to $+20$ degrees (Fig. 3.3(a)). Secondly, we perturbed the sensors either in a 1D arc around the head from -20 to 20 mm (Fig. 3.3(b) or within a 3D volume of $40 \times 40 \times 40$ mm^3 (see Appendix A section). For each perturbation, we then computed a forward model, estimated the most likely cortical current distribution, and obtained a Free Energy value.

(a) Orientation perturbation   (b) Position perturbation
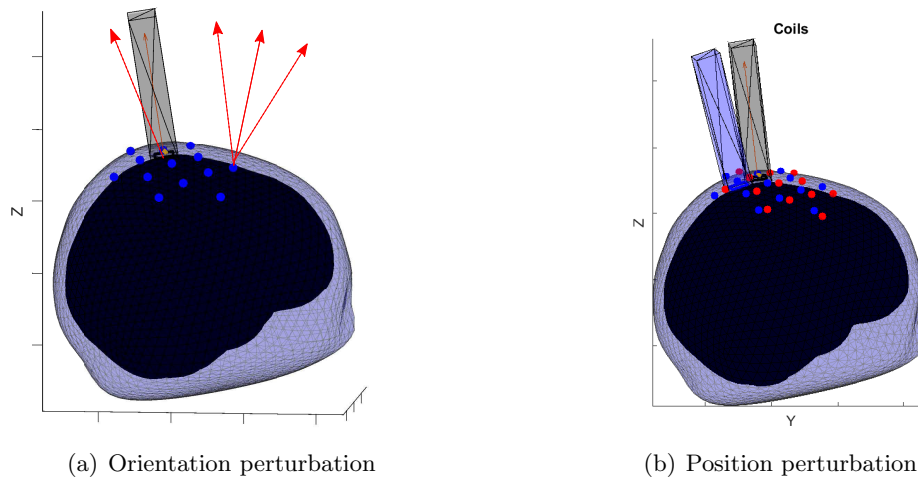
FIGURE 3.3: Perturbations of the sensor array. 3.3(a) Each sensor was independently perturbed from their true orientation (black) by a fixed angle in random direction (red). 3.3(b) The rigid sensor array was displaced from its true position with respect to the center of the brain (black) to new positions (red) within an arc spanning -20 to 20 mm (and subsequently a cube of $40 \times 40 \times 40$ mm^3)

Perturbation of the position and orientation: In order to perturb the geometry of the array, we used spatial transformation matrices. These matrices are used to represent the orientation and position of a coordinate system. In this case, the principal coordinates of the sensors are set in the MNI coordinates of the brain. As we know the coordinates of the sensors relative to the brain by the scanner-casts, we can compute the new coordinates of the sensors for each perturbation. Any combination of translation

and rotation of each sensor can be computed in a single 4 by 4 affine transformation matrix:

$$T = \begin{bmatrix} R & P \\ 0 & 1 \end{bmatrix}$$

where the sub-matrix $\mathbf{R} \in \mathbb{R}^{\mathbf{3 \times 3}}$ represents the orientation (or rotation) of the object in the space, and the vector $P \in \mathbb{R}^{3 \times 1}$ represents the position of the object in the space. An orientation error for any sensor can be computed by pre-multiplying the transformation matrix of the sensor with the transformation matrix of the perturbation (angle θ) in any axis:

$$\mathbf{R_x}(\theta) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & cos(\theta) & sin(\theta) & 0 \\ 0 & -sin(\theta) & cos(\theta) & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \mathbf{R_y}(\theta) = \begin{bmatrix} cos(\theta) & 0 & -sin(\theta) & 0 \\ 0 & 1 & 0 & 0 \\ sin(\theta) & 0 & cos(\theta) & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$\mathbf{R_z}(\theta) = \begin{bmatrix} cos(\theta) & sin(\theta) & 0 & 0 \\ -sin(\theta) & cos(\theta) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

A position error can be computed by rotating the whole array in any plane (here is rotated in the xy plane. We adapted the framework presented in [14] to explore across the perturbed arrays. Briefly, with each perturbed array a new gain matrix L_a (forward model) is computed using predominantly the single shell forward model [46], although we also made use of the single sphere model [59].

3.4   Results

Fig. 3.4 shows the averaged time courses from single channel OPM data (Fig. 3.4(a)), multi-channel data (Fig.3.4(a)) and SQUID data (with the 13 channels chosen for comparison highlighted in red. Those channels are the ones with the highest spatial correlation with the OPM array) (Fig. 3.4(c)); time zero corresponds to the median nerve stimulation impulse. The expected m20 evoked response is visible with the three experiments (red-dotted line). Note the scale changes in the axes with the OPM signals being 5-10 times larger in magnitude. As expected, the magnetic fields measured with the OPMs have a stronger response due to their proximity to the scalp; despite this

signal magnitude advantage however, the relative SNRs (at sensor level) are comparable across all three experiments (Fig.3.4(d)), presumably because much of the variability in the signal is of neural origin.

(a) Single channel OPM data

(b) Multichannel OPM data

(c) SQUIDs recordings

(d) SNR for the different modalities
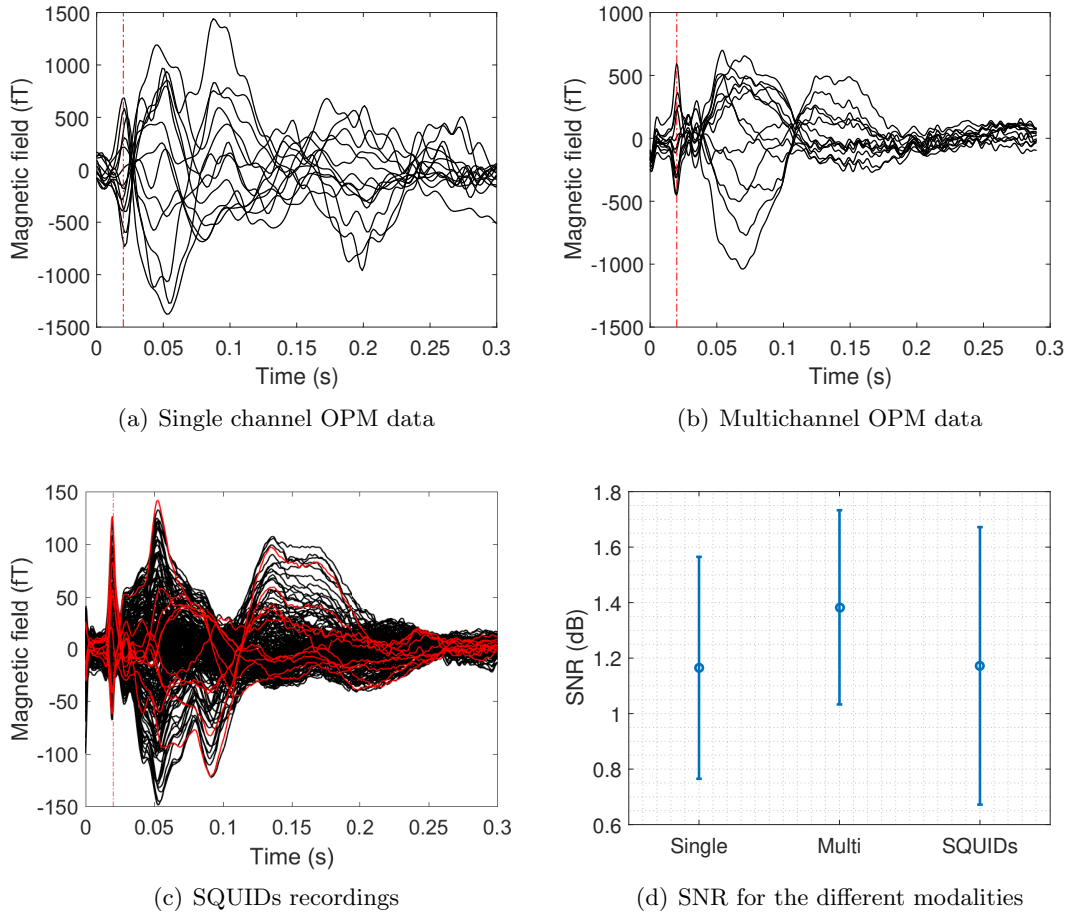
FIGURE 3.4: Empirical data. Sensor-level time courses for median nerve stimulation recorded with OPMs and SQUIDs. The evoked response at 20 ms (red dotted line) is highlighted. Black lines show time courses from a single channel, averaged over trials. 3.4(a) Single channel OPM used to sequentially record responses at 13 different locations across the scanner-cast, treated as a simultaneous measurement. 3.4(b) Multichannel OPM data recorded simultaneously with 13 different channels in the same scanner-cast. 3.4(c) SQUIDs recordings with the 13 channels that are spatially correlated to OPM channels highlighted in red (the 13 SQUIDs channels used are the ones with the highest spatial correlation with the OPM array). Note the scale changes between sensor types. 3.4(d) SNR calculated for the three experiments. Where $SNR = max_i(\frac{\sum_{pre} Y_i^2}{\sum_{post} Y_i^2})$, is the maximum ratio of post-stimulus to baseline power in any sensor.

3.4.1   Adding sensor orientation Error

In the first instance, we wanted to examine the sensitivity of our models to sensor orientation error and gain error. The logic being that sensitivity to error in the geometry is a prerequisite for any scheme seeking to optimise geometry. We also considered gain-error to account for other un-modelled sensor imperfections due to calibration or cross-talk issues. Individual sensor orientations were perturbed by orientation errors between -20 to 20 degrees in a random direction around their true orientation in steps of 2.5 degrees. A total of 30 **L** forward models were obtained for each orientation error (i.e. each forward model has all channels perturbed in a different random direction about their true axis by this amount). Additionally to orientation error we perturbed the models with gain errors of 5% and 20% (Fig. 3.5(a)).
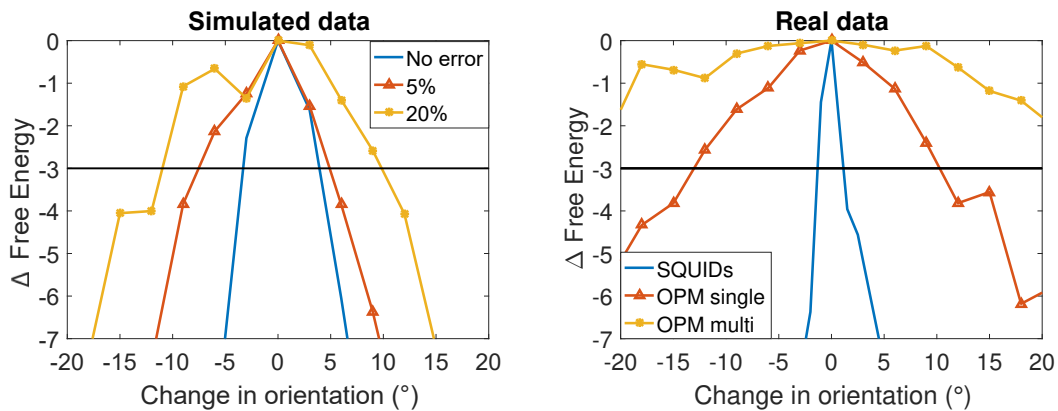
(a) Sensor Orientation error for simulated data   (b) (Software) Change in orientation for real data

FIGURE 3.5: 3.5(a) OPMs simulations. Sensitivity of model fit (Free Energy metric) to errors in sensor orientation: for perfect sensors (blue solid), sensors with gain errors of 5% (orange triangles), and gain errors of 20% (yellow circles). Adding gain error to the data results in broadening of posterior estimate on sensor orientation. Solid black line ($F = -3$) is the point at which the models become 20 times less likely than the best model. 3.5(b) Sensitivity of model fit to orientation errors added to real sensor recordings: for SQUID data (blue solid); single channel OPM data (orange triangles) and multi-channel OPM data (yellow triangles). Note that the sensitivity to individual channel orientation error is lower in the empirical OPM recordings than the SQUID system.

Fig. 3.5(a) shows the change in Free Energy as a function of channel orientation error for simulated data. The solid-line shows simulations with an idealized OPM sensor array. Note that the Free Energy peaks at zero error where the measured data can be most simply reconciled with the single generating source. For Free Energy values (on a log scale) -3 corresponds to models that are 20 times less likely. In the ideal sensor case we are therefore able to reject sensor geometries with more than ± 4 degrees

of intrinsic error as unlikely. Also shown are the effects of additional random gain-error (5%, 20%; red triangles and yellow circles respectively). These serve to blunt the orientation perturbation curves- adding 20% gain error to the sensors means that it is now only possible to confidently reject sensor geometries with greater than ± 12 degrees orientation error- although the most likely sensor geometry remains the true geometry.

Fig. 3.5(b) shows the same orientation perturbation curves but based around real measured data from the three MEG systems. All three datasets are also sensitive to perturbations of the geometry of the measurement sensors and suggest that the most likely orientation is the true one. The model used to describe the SQUID data is sensitive to orientation error of less than ± 5 degrees, the models used to describe the concatenated single channel OPM data being sensitive to orientation error of ± 12 degrees and the model used to describe the multi-channel OPM data is relatively insensitive to orientation error. We speculate that the difference between the single and multi-channel system OPM curves is that the concatenated single channel system is effectively a more homogenous system than the multi-channel system. The multi-channel system will suffer from sensor cross-talk and other factors (such as calibration, different intrinsic noise levels etc) of between-sensor variability. However, the difference between the SQUID and OPM curves ran counter to our expectation- which was that the OPM models would have the higher sensitivity to perturbation because of the marginally higher SNR [6, 8].

3.4.2   Movement of sensor array

The other scenario we considered was a rigid array of sensors of known relative geometry attached to the scalp. In this case the goal is to estimate the array location (as fixed whole) relative to the subjects anatomy. This could be locating a small array of OPMs strapped to the scalp surface or estimating the position and orientation of a generic helmet (e.g. bicycle helmet) containing the sensors. We begin by demonstrating the change in Free Energy as the sensor array is moved in an arc about its true position. Fig. 3.6(a) shows the effect of this movement on simulated OPM data. Again, the most likely array location is the true location, and the 95% confidence bounds on this location are ± 8 mm. As gain error is increased these error bounds become larger (i.e. the curves become wider).

Fig. 3.6(b) shows the same (software) displacement of the sensor array used to collect real OPM and SQUID data (i.e. error was added to the sensor array locations from a real data recording, and a search across a range of array locations was performed with the algorithm being agnostic to the true array location). Again we were encouraged to find that all forward models to explain these real data exhibit maximal model evidence at the

true array location, even though this location is unknown to the algorithm. We observe a similar pattern to the orientation error manipulations with real data, in that the models using data from SQUID sensors once again have the least positional uncertainty (± 2 mm); before reaching significance (i.e. very small changes in sensor array position are detrimental to the goodness-of-fit); followed by the OPM single channel data (± 3 mm) and the OPM multi-channel recordings (± 10 mm).
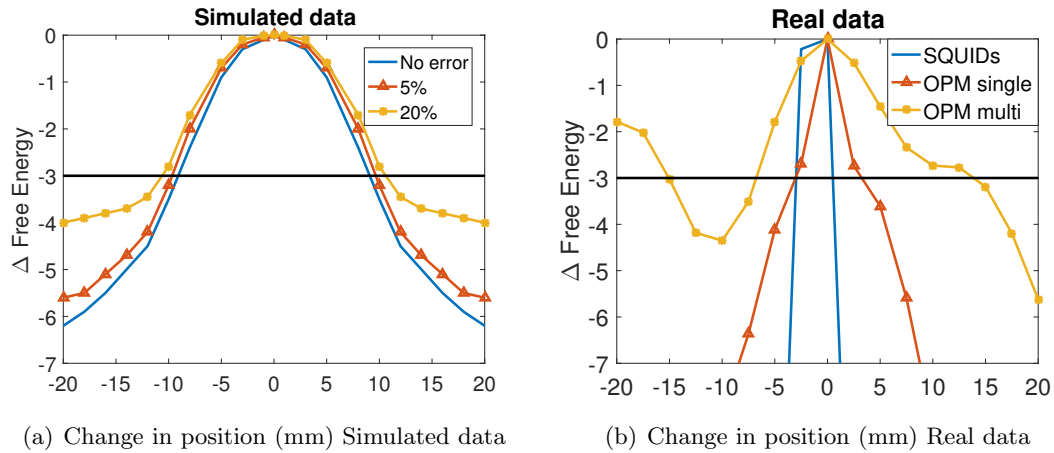
(a) Change in position (mm) Simulated data (b) Change in position (mm) Real data

FIGURE 3.6: Effect of sensor array displacement on goodness-of-fit, simulated and real data. 3.6(a) Simulated data. Sensitivity of Free Energy to errors in sensor array position: for perfect sensors (blue solid), sensors with gain errors of 5% (orange triangles), and gain errors of 20% (yellow circles). Adding gain error to the data results in subtle broadening of posterior estimate of the sensor array position. 3.6(b) Sensitivity of Free Energy to array position (ground truth based upon head and scanner-casts estimates) added to real sensor recordings: for SQUID data (blue solid); single channel OPM data (orange triangles) and multi-channel OPM data (yellow circles). Note that the Free Energy peaks at zero error where the measured data can be interpreted as the generating source.

3.4.3 Model Optimisation

In sections 3.4.1 and 3.4.2 we perturbed the geometry of the array in two ways. First, by perturbing the orientation of the sensors while the position remain fixed between -20 to 20°. Then, by perturbing the position of the rigid array in which the relation of the sensors remains the same through an arch of -20 to 20 mm around the xy axis. The obtained curves show that the true position is the one with the highest free energy. Such result suggests that we can find the true position of an array with only an approximate prior position. The practical problem is now to demonstrate how it is possible to locate a rigid sensor array with only approximate positional information based only on the field measurements, a volume conductor model and the cortical geometry. We did this by using the data from the single-channel OPM array in two ways. Firstly, using a simple 2D search passing over the known location of the sensor array. Secondly, by assuming

an initial uniform uncertainty over a 64 ($4 \times 4 \times 4$) cm^3 volume a-priori knowledge of sensor array location in any dimension.

3.4.4   Optimisation in one dimension

Assuming a rigid array we set a prior search space of $\xi = 40$ mm around the true position. Then, we used the Metropolis search algorithm detailed in the Appendix A. Four chains were simulated with single axis movement in which the algorithm had no information of the true array position respect to the brain, i.e. flat priors on location within [-20 20] mm of the true position. The Metropolis search was performed with 600 iterations per chain in four chains. Fig. 3.7(a) shows the change in the position of the array as it moved through an arc of 18 mm (i.e. the perturbed array started with an error of 18 mm). The initial value is represented with a green mark. Through each iteration of the Metropolis search (black dotted line) the position changes (via model comparison Fig. 3.7(b)) until convergence (blue mark). The error drops with each iteration (Fig. 3.7(c)) and after 250 iterations the algorithm oscillates near to the true position. Fig. 3.7(d) shows the prior and posterior estimate of the array location. The figure shows that the model estimate of the array position was 0.44 mm from our estimate of location based on the scanner-cast. The uncertainty (95 percentile) on this geometry estimate is also less than 1 mm (Fig. 3.7(d)).

(a) Orientation perturbation

(b) Position perturbation

(c) Position perturbation

(d) Position perturbation
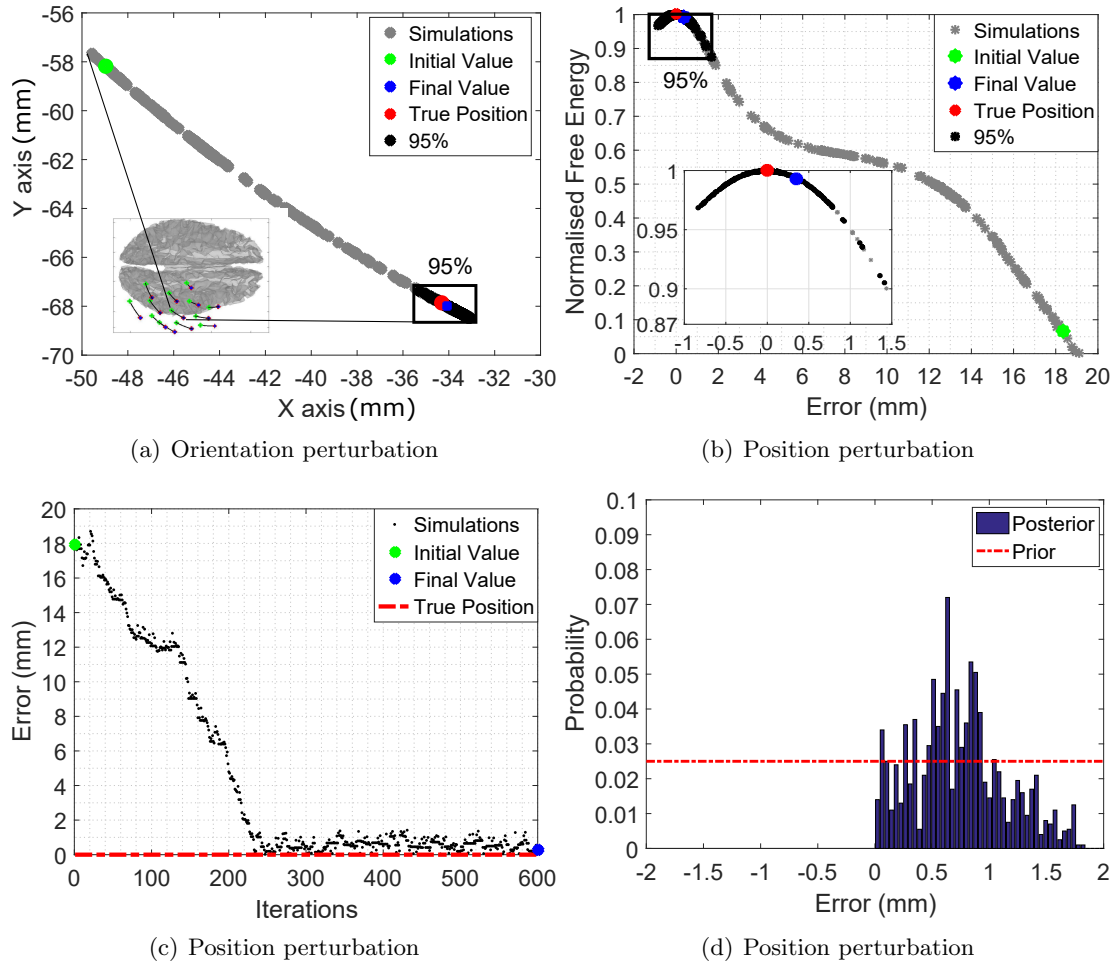
FIGURE 3.7: 3.7(a) Movement of the array through each iteration of the Metropolis search, the array moves through an arc in a 2D plane; the initial value is in 18 mm of error (green point), and evolves through each iteration (black points) until reach the final value (blue point). 3.7(b) Evolution of the Free Energy through each iteration. A first model is computed with the array centered of the sensors in the initial value (green point), then the inverse problem is solved and a Free Energy value is computed. The position of the array is updated through each iteration of the metropolis search until convergence (black points). The blue point represents the final position of the array while the red point represents the true position (as estimated from the scanner-cast). 3.7(c) Evolution of the distance error from the scanner-cast location, this error is unknown to the algorithm. 3.7(d) Prior and Posterior distributions of the array location (based on MEG data and uniform priors); zero represents the approximate array position on the scanner-cast.

3.4.5   Optimisation in three dimensions

Although the optimisation in 1D provides a clear illustration of the Metropolis process, it is not practically useful since positional uncertainty will rarely be constrained to lie in one dimension. To show how this method can be generalized to higher dimensional spaces we used the same Metropolis procedure but based on the assumption that sensor location was only known to within an approximate 3D volume of $4 \times 4 \times 4$ cm^3. Fig. 3.8 shows the prior cubic volume for the centre of the array (blue cube i.e the search space for the Metropolis algorithm); alongside the posterior confidence interval (black ellipsoid i.e the estimation of the centre of the array with the BMA procedure) and the scanner-cast estimate of this sensor location (red i.e the measured centre of the array via scanner-cast). After the optimisation, the Metropolis search and BMA estimated the posterior mean array position to be 4 mm displaced from that we expected from the scanner cast. The posterior confidence volume on this location was 0.1019 cm^3 i.e. a 600-fold reduction on the prior volume.
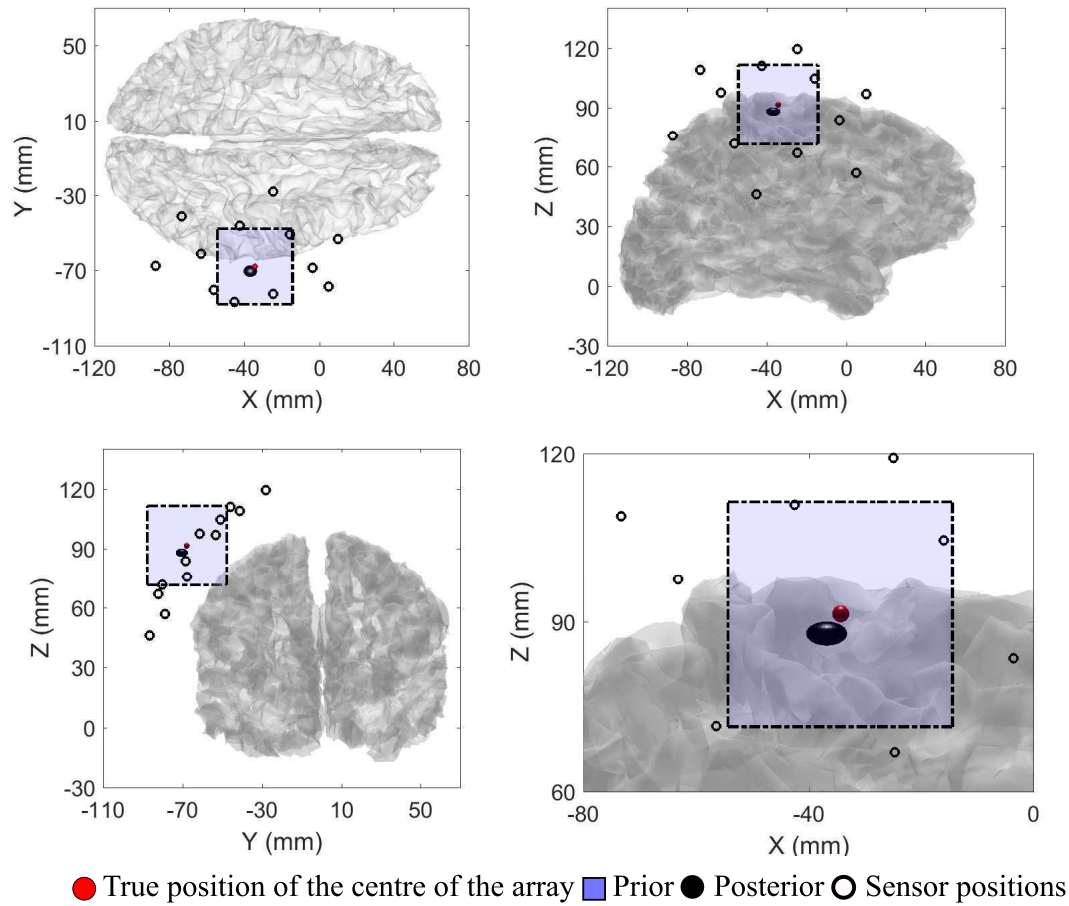
FIGURE 3.8: Optimisation in three dimensions (sensor space). The dotted cube shows the original $4 \times 4 \times 4$ cm^3 uncertainty on array location. The 95% confidence ellipsoid (black) shows the posterior location of the central sensor (and hence the whole rigid array). The location of the central sensor based on the scanner cast information is show as a red dot. Lower right panel is a magnified sagittal view.

It is also possible to view the consequence of the refinement of sensor position at the source level. Fig. 3.9 represents the search space in the source space, i.e., the areas that could be estimated if the geometry of the array is not found with precision. Estimating the source level activity based on our prior knowledge of sensor position $(4 \times 4 \times 4$ cm$^3)$, gives a distribution of (of peak locations) than can be described by the 95% confidence ellipsoid (blue) in Fig. 3.9 With the BMA step we are able to pool estimates from across a range of optimisation steps and weight them by their model evidence. Also shown in Fig. 3.9 is the posterior confidence volume on the peak source location after the Bayesian Model Averaging over sensor geometries (black). The model optimisation reduces the confidence volume on peak location from 34.90 cm^3 to 0.05654 cm^3. The centre of the optimised confidence volume is 5 mm from the source estimate when using the scanner-cast location as ground truth (red dot). The BMA step gives a degree of robustness to the process and importantly provides us useful posterior estimates of the head location and an estimation of current distribution with a confidence interval.
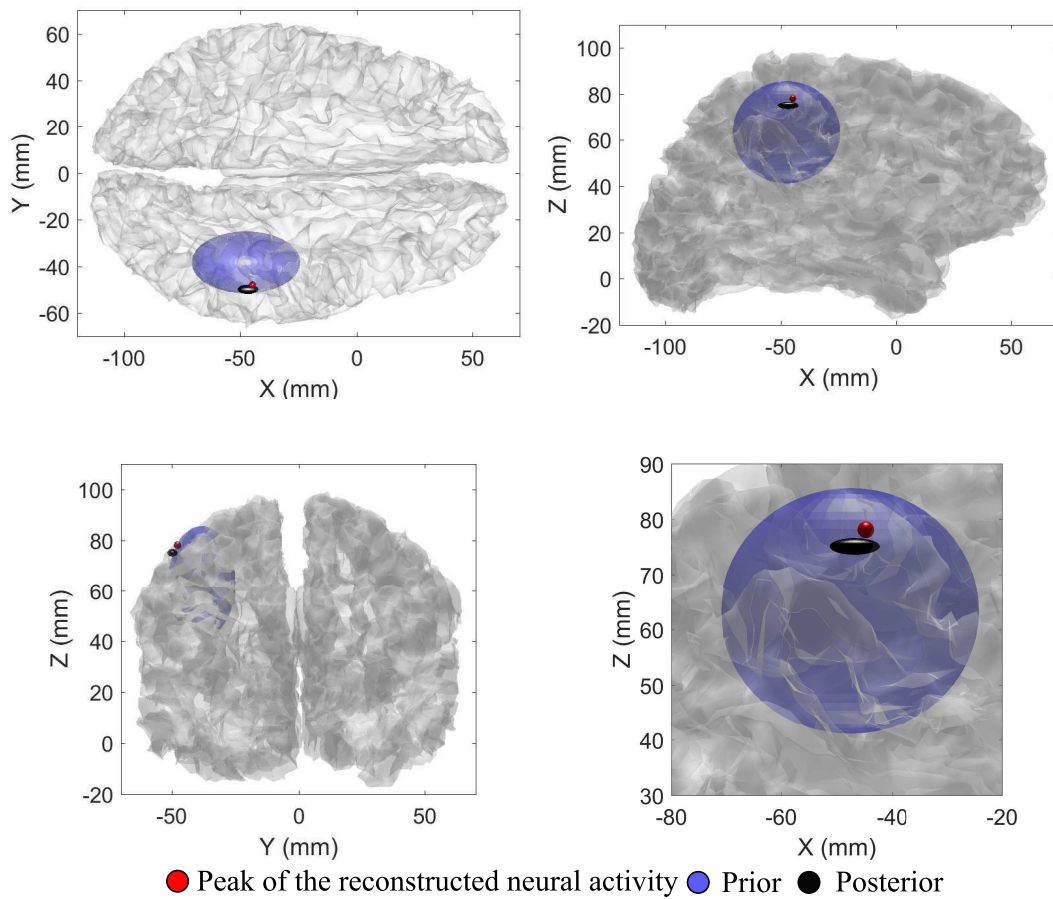
FIGURE 3.9: Optimisation in three dimensions (source space). Source estimates with confidence volumes shown in three orthogonal views. The red sphere represents the peak of the reconstructed neural activity when reconstructed with sensors at the scannercast locations. The initial sensor uncertainty gives rise to a prior distribution on the peak of the electrical activity (blue ellipsoid; based on reconstructions over 30 sensor locations distributed randomly across the prior volume). The black ellipsoid is the posterior estimate of electrical activity after BMA. The estimated source location when the sensor array location is unknown is 5 mm from the peak source location as estimated using the scannercast information.

3.5 Summary and Discussion

Here, we showed how model evidence is a useful metric to judge not only the quality of the source reconstruction [19, 54], but also the quality of the forward model [53, 60]. Model evidence is however data-dependent and cannot be compared across datasets or MEG systems. Here, we introduced the idea of quantifying how sensitive a given system is to geometrical perturbation. We demonstrated how the width of these precision curves could be used to compare different MEG systems or MEG system architectures.

We were initially surprised by our finding that our models of the SQUID system were less tolerant to errors than the OPMs. Especially given the simulation work of [6] who

showed that we would expect an idealized OPM system to have higher SNR (due to scalp proximity) and therefore more sensitivity to perturbations in the lead-field. We speculate why this might be the case in our data: firstly, in our recordings, the actual measurement SNRs were only marginally higher for OPMs than SQUIDs- despite the scalp proximity (Fig. 3.4). Whilst we might expect an increase in the white noise level by a factor of two for the OPMs, the OPM signal was around 5-10 times greater. We speculate that the limiting noise floor in the recording of the median nerve SEF is maybe a feature of the cortical response rather than the instrumentation. Additional sources of noise in the OPM measurements are that the OPMs are currently calibrated sequentially (and for each recording session) which we estimate adds around $5 - 10\%$ noise to the true OPM gain.

We have demonstrated how the spatial parameters (position and orientation) of a sensor array can be physically characterized based on magnetic fields derived from the human brain. In addition to removing the dependence on a scanner-cast, we can also dispense with traditional co-registration procedures and the associated subjective identification of arbitrary scalp landmarks. The co-registration here is performed with respect to inner skull anatomy (cortex and inner skull boundary) and unlike typical co-registration procedures (but see [14]), the geometrical uncertainty is directly factored into the source estimate giving realistic confidence bounds. For example, the experimenter needs only to specify that the array is approximately above the right ear (with a 64 cm^3 volume) and the algorithm is able to reduce this uncertainty by 600-fold to 0.1019 cm^3.

Here we used only a 3 parameter optimisation of a fixed array but the algorithm directly generalizes to optimisation over much larger parameter spaces (for example when only the topology of the array is known). The main consideration being the additional amount of data required. Importantly, as OPM devices are becoming wearable, we can expect subjects to tolerate the scanning environment for considerably longer periods, we will likely have far more data available with which to perform such optimisations. This would mean that the physical characterization of the sensor array and optimisation of forward models could be performed on data orthogonal to that under-scrutiny. For example, using stationary parcellations of resting-state data [21]. Additionally, we have measurements of magnetic fields tangential (rather than radial) to the cortical surface that we still have not used [58].

Although all of the data were collected from the same individual wearing either scanner or head cast, there were however differences in the recording paradigms. First, the SQUID data were collected based on right rather than left median nerve stimulation. Secondly, the Inter stimulus interval for the multichannel OPM and SQUID measurements was 0.5 s, in contrast to 1.9 s for the single channel OPM data which we know

will influence the evoked response components profile[61]. We, therefore, cannot rule out that there is some disparity in how well the data are modelled at the source level, which could in turn change the steepness of the geometrical tuning curves. We also tested the possibility that the SQUID tuning curves to position and orientation might benefit from the 5 cm baseline axial gradiometer configuration, but found negligible theoretical diference.

The problem of uncertain sensor placement is not specific to OPM MEG. In [62] have shown that inaccuracies of EEG electrode coordinates form an error term in the forward model and ultimately in the source reconstruction performance. This error arises from the combination of both intrinsic measurement noise of the digitization device and manual coregistration error when selecting fiducials on anatomical MRI volumes. OPMs pose additional challenges over EEG in that neither orientation nor position will be known in a more flexible setup. These problems will be yet more acute for the OPMs because the sensitivity to modelling errors is highly dependent on SNR [6, 8].

In this study, we have approximated the OPM as a point measurement system. In reality, the volume of the gas exposed to the laser light has maximal dimension of 3 mm. This distance is relatively large given that the OPM sensors may now sit < 20 mm from the brain. The addition of appropriate integration points within this volume would be a useful avenue for further study.

We made use of the scanner-cast here in order to provide some ground-truth on sensor position and orientation. However, some skew in the position of the cast on the head is possible (we estimated this to be around ± 3 mm, ± 5 degrees). We do not know therefore whether to attribute the final discrepancy (4 mm) between scanner-cast measurements and algorithm estimates position to the cast or the algorithm. But we note that the algorithm gives us posterior confidence bounds on the array location of better the 0.1019 cm^3. We see one use of this algorithm is to further refine our geometrical estimates from the scanner-cast.

Chapter 4

Enhanced data covariance estimation

4.1 Introduction

Estimating brain activity with magneto/electroencephalography (MEG/EEG) has been increasingly employed as a non-invasive technique for understanding the brain functions and neural dynamics. However, one of the main open problems when dealing with MEG/EEG data is its non-Gaussian and non-stationary structure. In this chapter, based on Multiple Kernel Learning (*MKL*), we combine a set of mapping functions equipped with distinct bandwidth sizes to encode different notions of similarity, intending to extract more accurately the predominant complex dynamics hidden in MEG/EEG sensors. To highlight the prominent data relationships, the MKL methodology employs a weighted mixture of Gaussian kernels (termed WM-MK), enhancing the data-driven covariance estimation and resulting in a more reliable source reconstruction. Obtained results of validation on simulated and real-world MEG/EEG data demonstrate to what extent a proper kernel combination enables increasing the performance of well-known distributed solutions to the MEG/EEG inverse problem.

4.2 Multiple Kernel Learning

Instead of linearly computing the covariances, the kernel-based estimation accounts for the non-stationary behavior of available data, encoding the relationship between each couple of *M/EEG* channels $\{\boldsymbol{y}_c, \boldsymbol{y}_{c'}\}$ into the following kernel function [63]:

$$\kappa(\boldsymbol{y}_c, \boldsymbol{y}_{c'}) = \langle \phi(\boldsymbol{y}_c), \phi(\boldsymbol{y}_{c'}) \rangle, \quad \forall c, c' \in N_c \tag{4.1}$$

where $\boldsymbol{y}_c \in \mathbb{R}^{N_t}$ is the time-course of c−th MEG/EEG channel and $\phi(\cdot){:}\mathbb{R}^{N_t}{\to}\mathcal{H}$ maps from the original domain, $\mathbb{R}^{N_t}$, into a Reproduced Kernel Hilbert Space (*RKHS*), $\mathcal{H}$. Notation $\langle\cdot,\cdot\rangle$ stands for the inner product.

The *RKHS* dimension tends to infinity (i.e., $|\mathcal{H}|{\to}\infty$), so that $|\mathbb{R}^{N_t}|{\ll}|\mathcal{H}|$, $\phi(\cdot)$ cannot be directly calculated. Instead, we use the well-known *kernel trick* to compute Eq. 4.1 through a distance-based positive definite and infinitely divisible kernel function, $k_{c,c'}{=}\kappa\left(\mathrm{d}(\boldsymbol{y}_c, \boldsymbol{y}_{c'})\right)$, provided a distance operator $\mathrm{d}{:}\mathbb{R}^{N_t\times 1}{\times}\mathbb{R}^{N_t\times 1}{\mapsto}\mathbb{R}^+$. Hence, applying κ over each channel-pair produces a kernel covariance estimator $\mathbf{K}\in\mathbb{R}^{N_c\times N_c}$ in the projected *RKHS* space. Here, the kernel functions are Gaussian:

$$\kappa(\boldsymbol{y}_c, \boldsymbol{y}_{c'}, \sigma) = \exp(-||\boldsymbol{y}_c - \boldsymbol{y}_{c'}||_2^2/\sigma^2) \qquad (4.2)$$

where $\|\cdot\|_2^2$ is the squared L_2 norm.

To capture the meaningful data relationships hidden in the input space, the bandwidth $\sigma\in\mathbb{R}^+$ is fixed by a *MKL* method that linearly combines N_m pre-defined kernel functions, each one equipped with different bandwidth, as follows:

$$\mathbf{K} = \sum_{m\in N_m} \alpha_m \mathbf{K}_m, \qquad (4.3)$$

where each term $\mathbf{K}_m\in\mathbb{R}^{N_c\times N_c}$ is a potential non-linear kernel to be combined, being $\alpha_m\in\mathbb{R}^+$ its corresponding weighting parameter.

For calculating the linear parameter combination, we follow an information-theoretic approach based on the Kullback-Leibler (*KL*) divergence (noted as $\mathrm{d}_{KL}(\cdot,\cdot)$) between a mixture of kernels $\mathbf{K}$ and the input kernel matrix $\mathbf{K}_Y\in\mathbb{R}^{N_c\times N_c}$, having conic sum constraints upon the weighting parameters [29]:

$$\min_{\alpha_1,\ldots,\alpha_m} \{\mathrm{d}_{KL}(\mathcal{N}(\mathbf{0}, \mathbf{K}), \mathcal{N}(\mathbf{0}, \mathbf{K}_Y))\} \qquad (4.4)$$
$$\mathrm{s.\,t.}: \sum_{m\in N_m} \alpha_m = 1, \quad \alpha_m > 0.$$

Optimizing the constrained formulation in Eq. 4.4 enables extracting the relevance weight (prominence) of each kernel employing conventional methods, such as the projected gradient descent [64]. As a result, the combined kernel $\mathbf{K}$ is obtained using Eq. 4.3.

4.3 Experimental Setup

To capture a broader class of data dynamics than the variability measured by a single-kernel bandwidth, the introduced *WM-MK* approach uses a set of N_K Gaussian kernels with different bandwidth to feed the whole multi-kernel estimator within the expanded variability span.

For testing, the evaluated MEG/EEG brain imaging methods are implemented on SPM12 (http://www.fil.ion.ucl.ac.uk/spm). Note that for validating the simulated databases, the reconstruction is carried out using the same head model for which the data is produced. Although the use of this procedure commits the so-called "*inverse crime*",which is a biasing effect occurring when the lead field matrix used to generate the synthetic data is the same for estimating the reconstructed brain activity, leading to overly optimistic results [65], it still allows validating the performance of these methods [17]. Fig. 4.1 illustrates the proposed approach.
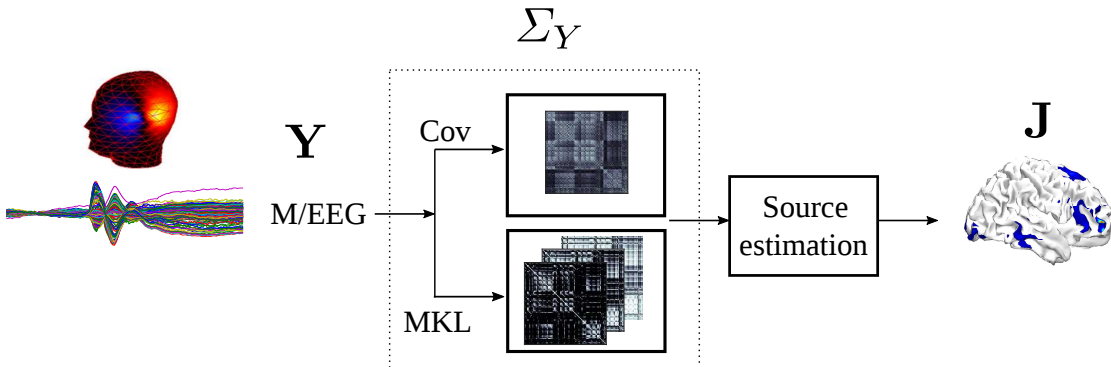
FIGURE 4.1: Benchmark of compared single kernel and multikernel data covariance estimation for ESI.

During validation, the following datasets are tested: *i*) two collections differently simulated of *MEG/EEG* data, incorporating Gaussian and non-Gaussian noise conditions. *ii*) A well-suited real MEG/EEG dataset for which the ground truth is the activity elicited by a multi-modal study, comprising EEG and MEG data.

4.3.1 Simulated MEG/EEG data

Intending to generate realistic M/EEG data, the simulated recordings must be designed using the minimal realistic case of brain activity simulation described in [66], holding the following guidelines: (i) A realistic forward model, (ii) Realistic source locations, being confined to the cortical manifold, (iii) Variable locations, different spatial extents, and depths of the sources, (iv) The presence of independent background brain processes with

1/f (pink noise) spectra, and (vi) The presence of measurement noise at different SNR levels. With this aim, the following two simulation benchmark datasets are employed:

SDB1: This dataset is generated for testing the compared responses to interacting brain activity with stationary temporal dynamics. In this case, two active sources are modeled as a stationary band-limited linear auto-regressive process, each one randomly allocated in one of the eight brain octants or Regions of Interest (ROIs) [66]. Besides, 500 background uncorrelated sources, created as pink noise generators, are placed at random locations of the cortical mesh. The simulation set employs the lead-field values precomputed for a realistic head model (termed *New York Head* [67]), having a detailed segmentation of six tissue types (scalp, skull, cerebro-spinal fluid, gray matter, white matter, and air cavities) and using an MRI resolution of 0.5 mm^3. Additionally, the number of electrodes and cortical locations are fixed to N_c=108 and N_d=2000, respectively. Thus, we perform 100 runs, each one lasting N_t=10 s and sampled at 200 Hz with a randomly selected Signal-to-Noise Ratio (SNR).

SDB2: This dataset is designed to test the compared methods under non-stationary brain activity and non-Gaussian noise conditions. With this purpose, a single-trial MEG dataset is reproduced for N_c=270 sensors using the Montreal Neurological Institute (MNI) template implemented in SPM12. The source time-courses are generated with N_t=1 s, and sampled at 160 Hz. Further, we create the time-locked brain activity that involves a couple of active sources by the real part of a Morlet wavelet, which is commonly accepted in studies of evoked related potentials (ERP) [68]. In particular, the central frequency of each wavelet is randomly sampled from a Gaussian distribution with mean 9 Hz and standard deviation 2 Hz. Then, the simulated source activity is mapped to MEG sensors using a tessellated surface in the gray-white matter interface, fixing N_d=8196 vertices (number of possible source locations) with the source orientations fixed orthogonally to the cortical surface. In this case, the lead fields are computed employing the Single-Shell volume conductor model of SPM12. This procedure is also conducted using non-Gaussian noise modeled with values: skewness s=0.3 and kurtosis k=4. As a result, a number of 50 trials are generated with varying noise conditions, adding random white noise to the data with a different SNR: $\{-5, 0, 5, 10, 15, 20\}$ *dB*.

4.3.2 Real-world MEG/EEG data

The multi-modal human neuroimaging dataset [69], noted as **RDB1**, holds EEG, MEG, as well as functional and structural MRI data acquired from N_S=16 subjects [69], and it is intended to provide a comparison benchmark for visual stimulation. During acquisition, each subject undertook multiple runs of a simple task performed on a large

number of faces labeled as follows: Famous, Unfamiliar, and Scrambled. The MEG data hold N_c=306 channels of the Elekta VectorView system employed to simultaneously collect the EEG recordings from N_c=70 electrodes (using the nose reference). Six sessions, each one lasting approximately ten minutes, were acquired from each subject, while they judged the type of stimulus presented (face or scrambled face), providing nearly 300 trials in total for each one of three testing conditions. The trials, belonging to the same condition were averaged to perform a single ERP per stimulus. The main objective of this dataset is to analyze the related amplitude of N170 component, which is a component of the *ERP* that reflects the neural processing of faces.

4.3.3 Tested source reconstruction approaches

We compare performance of the MEG/EEG brain imaging methods Loreta, Empirical bayesian beamformer, and greedy search (LOR, EBB, and GS) regarding their ability to accurately estimate the source location and to reconstruct the time-courses from the simulated time-locked activity. For assessing the influence of enhanced covariance estimation on the tested inversion methods, the following estimation approaches are compared: *i*) Sample Covariance matrix computed in Eq. 2.7 (noted as *COV*) as a baseline linear estimation widely used in practice. *ii*) A single Gaussian kernel with a bandwidth calculated from the median of data distances (*MED-SK*), which is a nonlinear estimate that provides a significantly larger bandwidth $\sigma \in \mathbb{R}^+$ for globally encoding the data relationships [63]. *iii*) A single Gaussian kernel with a bandwidth computed by a more elaborated estimate (using information metrics, *IM-SK*), seeking for higher separability among data groups [70]. *iv*) The proposed weighted mixture of Gaussian kernels using MLK (*WM-MK*).

In order to capture a broader class of MEG/EEG data dynamics than the variability measured by a one-kernel bandwidth, the introduced *WM-MK* approach uses a set of N_K Gaussian kernels with different bandwidths to feed the whole multi-kernel estimator within an expanded variability span. In this sense, we set N_K=30, where the bandwidths create a mesh ranging between $0.1\sigma_0$ and $10\sigma_0$, being $\sigma_0 \in \mathbb{R}^+$ a seed value calculated as in either approach: *MED-SK* or *IM-SK*. Furthermore, the input kernel matrix needed to learn the linear combination weights is computed as recommended in [29]:

$$\mathbf{K}_y = \frac{1}{N_t}\mathbf{Y}\mathbf{Y}^\top.$$

4.3.4 Evaluation metrics

The source reconstruction performance is assessed on the temporal and spatial domains using the following metrics:

- *ROI–localization*: $r_s \in \mathbb{R}[0, 1]$, this measure looks for the true octants, comprising the simulated sources against their estimated values. If both octants coincide, $r_s = 1$. Otherwise, $r_s = 0$. Consequently, the higher the r_s, the better the *ROI* localization.

- *Spatial accuracy index*: $d_s \in \mathbb{R}[0, 1]$, it computes the Euclidean distance between the actual simulated location and the position of the source with the highest power, estimated as the most powerful value of the computed primary current density matrix. Thus, the lower the d_s, the better the source reconstruction.

- *Temporal accuracy index*: $\epsilon_t \in \mathbb{R}[0, 1]$, computed as the average of maximum dipole correlation over the simulated sources, it measures the shape similarity between simulated and time-courses extracted from dipoles. The dipole correlation is estimated through the maximum value of the Pearson Correlation Coefficient across the dipoles related to each simulated source.

- *Earth-movers distance*: $\rho_s \in \mathbb{R}^+$, it estimates the spatial distribution of dipole-wise power, employing the rate between the neural activity and true source power. ρ_s measures the needed effort to transform the estimated power distribution into the actual distribution by transporting the probability mass. Thus, lower ρ_s values correspond to a better reconstruction.

- *Kernel alignment*: $\kappa_s \in \mathbb{R}[0, 1]$. This empirical alignment, estimating the similarity between two matrices, is calculated as follows:

$$\kappa(\mathbf{K}_i, \mathbf{K}_j) = \frac{\|\mathbf{K}_i, \mathbf{K}_j\|_F}{\sqrt{\|\mathbf{K}_i, \mathbf{K}_i\|_F \|\mathbf{K}_j, \mathbf{K}_j\|_F}} \tag{4.5}$$

  where $\|\mathbf{K}_i, \mathbf{K}_j\|_F$ is the pairwise Frobenius inner matrix product. Consequently, the closer to one the κ_s value, the more alike the matched matrix kernels. Therefore, rather than evaluating the source reconstruction, κ_s assesses the brain response similarity as the averaged pairwise alignment between subjects.

During testing of source reconstruction methods, a separate performance measure is used following the purpose of each validated dataset. Namely, in the case of **SDB1**, r_s and ϵ_t are employed to estimate the actual ROI set (octants), comprising brain activity and time-courses. For **SDB2**, other measures (d_s and ϵ_t) compute the actual location

and time-course of sources. Regarding **RDB1**, the ρ_s value measures to what extent the estimated reconstruction is close to the available ground truth, while κ_s estimates the brain activity homogeneity over the sensor and source spaces across subjects presented with a standard stimulus. Besides, the influence of the examined covariance estimators is appraised for all datasets.

4.4 Results

4.4.1 *SDB1* benchmarking results

Table 4.1 presents the results obtained by each comparing method in terms of *ROI*-localization performance r_s, and averaged temporal accuracy index ϵ_t, obtained after 100 trials. Because of the imposed realistic constraints upon the source covariance matrix, GS achieves the best source localization accuracy, outperforming EBB and LOR, respectively. This result is expected based on previous studies [54]. Note that for r_s, the standard deviations are not provided since each estimate accounts for the percentage of times that both octants are correctly identified across all trials. In this sense, r_s does not reflect the case where only one ROI where correctly estimated. As EBB suffers with correlated sources, it is expected that EBB reflects lower results compared with GS.

TABLE 4.1: Accuracy of the *ROI* selection and correlation of reconstructed time-courses computed on **SDB1**. The values marked in bold are the best performance per row. Notation **Av** stands for average.

ESI	*COV*	*IM-SK*	*MED-SK*	*WM-MK*	**Av**
		ROI localization index, r_s			
LOR	54	66	62	**71.0**	63.25
EBB	73	74	76	**77.0**	75.00
GS	85	85	87	**88.0**	86.25
Av	70.6	75	75	**78.6**	
		Temporal performance index, ϵ_t			
LOR	53.31 ± 6.24	61.22 ± 5.31	63.65 ± 4.87	**68.13 ± 4.12**	61.57
EBB	84.98 ± 3.65	86.24 ± 2.95	85.37 ± 3.08	**87.09 ± 2.75**	85.92
GS	89.43 ± 2.78	88.63 ± 1.96	90.25 ± 2.05	**92.36 ± 1.68**	90.16
Av	75.9	78.7	79.7	**81.7**	

Regarding the data covariance, the linear *COV* estimate provides the lowest accuracy, on average, in comparison to the kernel-based methods. Thus, either single kernel approach reaches a comparable source reconstruction accuracy. Nonetheless, the best accuracy is achieved by the multi-kernel approach *WM-MK* over both space and time domains. Additional evaluation is carried out to assess the robustness of *WM-MK* estimator under variations of three essential parameters:

i) *Kernel bandwidth initialization.* Two approaches were implemented to calculate the seed kernel bandwidth σ_0, *IM-SK* and *MED-SK*. However, the accuracy achieved by *WM-MK* with either of them was similar, with $\rho_s = 87$ and $\epsilon_t = 91.72 \pm 1.48$ for *WM-MK*(IM), and $\rho_s = 89$ and $\epsilon_t = 92.36 \pm 1.68$ for *WM-MK*(MED).

ii) *Variability span.* The achieved *ROI*-localization accuracy is analyzed within a couple of test ranges, which are fixed as to capture enough information from input data, applying the seeds computed by the median and IT approaches. Fig. 4.2(a) shows that the value $N_K \geq 30$ is enough to implement the multi-kernel representation for both tested ranges of σ: $0.1\sigma_0$–$3\sigma_0$ and $0.1\sigma_0$–$10\sigma_0$, yet, the expanded interval results in a better source reconstruction accuracy.

iii) *Number of combined kernels N_K* that indirectly adjusts the kernel size set, encoding to a different extent the complex data relationship and thus better representing the complexity of MEG/EEG dynamics, as seen in Fig. 4.2(b).

(a) Tunning of N_K (b) Estimated values of σ
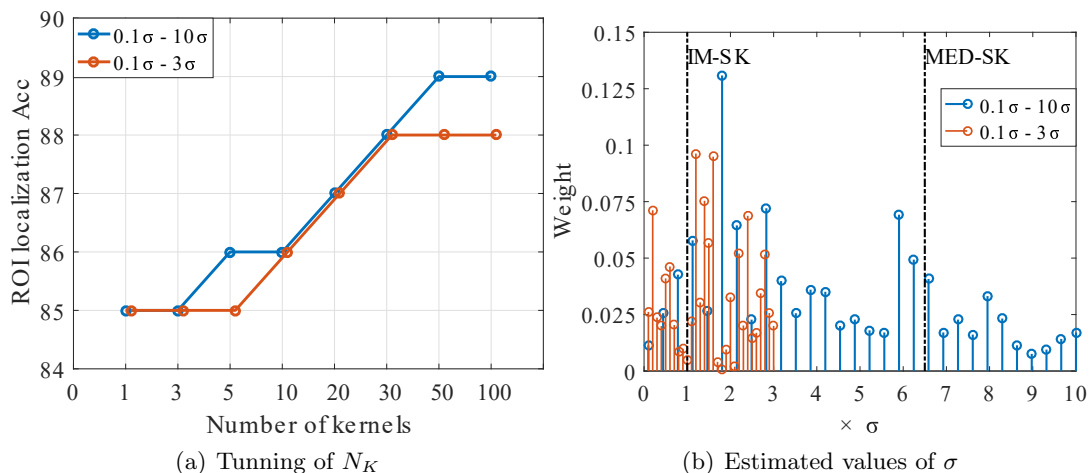
FIGURE 4.2: 4.2(a) Influence of the number of considered Gaussian Kernels N_K on the source reconstruction for both different seed values: IM-SK and MED-SK. 4.2(b) Estimated weights for a value fixed to N_K=30. with two different variablity spans.

4.4.2 Performed source reconstruction on *SDB2*

Fig.4.3 shows that all compared brain imaging methods increase d_s as SNR raises, regardless of the tested covariance estimator or either explored measurement noise model (Gaussian and non-Gaussian). Additionally, Fig.4.3(a) reveals that the prior simplicity of LOR makes it the most affected by the measurement noise, while the *EBB* solution provides an accuracy close to *GS* (see Figs. 4.3(b) and 4.3(c)), for which the accuracy d_s improves up to 5 mm compared with the LOR reconstruction. EBB achieves even lower d_s values having non-Gaussian noise (see bottom row). When comparing the covariance

estimation approaches, *COV* gives the worst spatial accuracy in almost all cases, being outperformed by *IM-SK* and *MED-SK*, which perform similarly. *WM-MK* provides the best d_s value, regardless of the involved method. Note that these results hold for MEG data corrupted by Gaussian or non-Gaussian noise, though the latter caused higher degradation.

The intuition behind the spatial accuracy index d_s is the source localisation error, i.e., the Euclidean distance between the simulated source and estimated sources. Thus, the lower the d_s, the closest the source reconstruction with the simulated one. In this sense, the results obtained with the simulations with Gaussian and non-Gaussian noise suggests that the EBB WM-MK provides the best source reconstruction in a theoretical real scenario (the 0 dB SNR) and in the worst scenario with non-Gaussian noise.
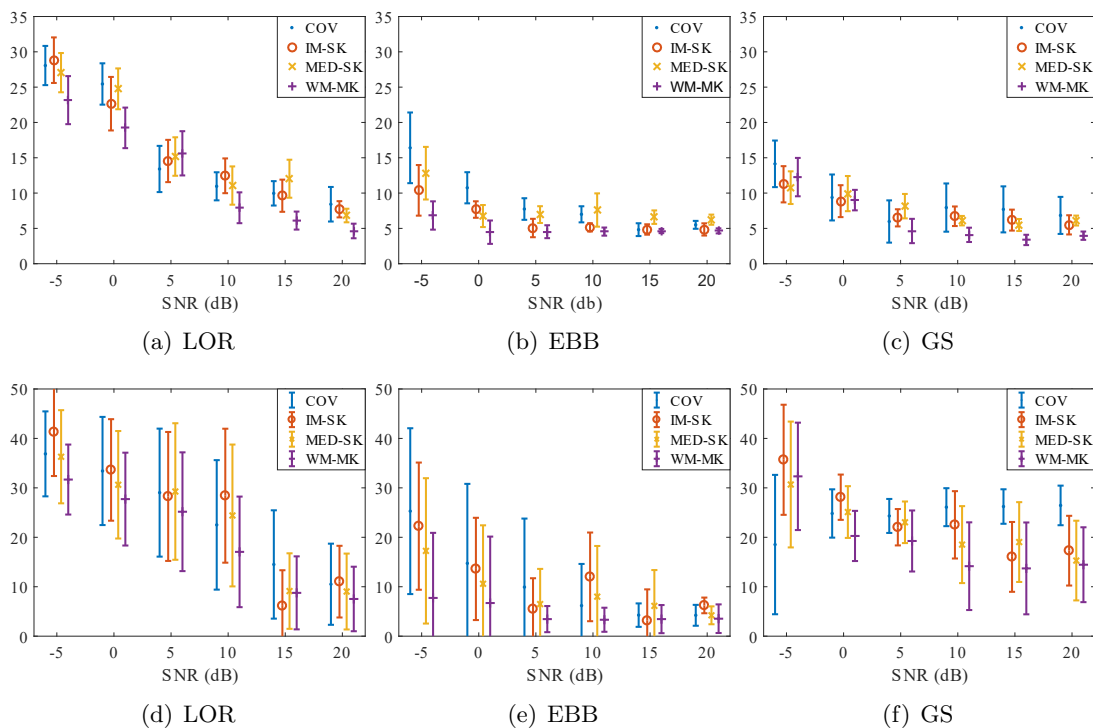
FIGURE 4.3: Spatial accuracy index d_s achieved by tested methods with different covariance approaches. Top and bottom rows depict achieved results under Gaussian and non-Gaussian noise conditions, respectively.

Furthermore, we explore the confidence interval of the covariance estimations, using the GS solution as the one producing the best spatial reconstruction accuracy. To this end, a paired t-test is employed, assuming as null hypothesis that there are no significant differences between each pairwise tested estimators concerning the d_s values (averaged across 50 trials). Otherwise, the alternative hypothesis affirms that one of the mean d_s values is confidently lower. Although all kernel-based covariance estimations outperform the linear *COV* method, the proposed WM-MK estimation yields the best confidence

TABLE 4.2: Paired *t*-test comparing the tested covariance estimations with the spatial accuracy index d_s. In the *t*-test alternative hypothesis, the method named in the right is assumed to have a significantly lower d_s mean value.

SNR	−5			0			5		
	h	t	p	h	t	p	h	t	p
Cov vs. IM-SK	1	4.7304	0	0	−1.629	0.1095	1	3.3441	0.0016∗
Cov vs. MED-SK	1	5.4485	0	1	−2.610	0.0120∗	0	−0.7086	0.4819
Cov vs. WM-MK	1	3.1384	0.0029 ∗∗	0	−0.161	0.8724	1	4.4935	0 ∗∗
IM-SK vs. MED-SK	0	0.0648	0.9486	0	−1.487	0.1433	1	−4.8435	0 ∗∗
IM-SK vs. WM-MK	1	−2.7642	0.0080 ∗∗	0	1.857	0.0693	1	3.7174	0 ∗∗
MED-SK vs. WM-MK	1	−3.1465	0.0028 ∗∗	1	3.366	0.0015∗	1	7.7800	0 ∗∗
Cov vs. IM-SK	0	−0.3161	0.7542	1	10.552	0	1	16.6182	0
Cov vs. MED-SK	0	−1.6015	0.1201	1	7.443	0	1	12.0668	0
Cov vs. WM-MK	1	−3.9143	0.0005	1	6.851	0	1	5.4172	0
IM-SK vs. MED-SK	1	−9.4138	0	1	−18.391	0	0	−21.3205	0
IM-SK vs. WM-MK	1	3.6292	0.0011	0	−1.137	0.2645	0	−1.1252	0.2697
MED-SK vs. WM-MK	0	1.7996	0.0823	0	0.575	0.3407	0	0.9688	0.3407

interval, demonstrating that the null hypothesis is rejected at a significance level $p=0.01$ over a wide range of tested SNR levels with either noise model: Gaussian (see top row in Tab. 4.2) or non-Gaussian (bottom row).

We also present a visual inspection of the spatial accuracy in Fig. 4.4. It displays in the top row three views of two simulated sources of auditory and frontal cortices, adding Gaussian noise at SNR=5 *dB*. Coincident with our previous results, the LOR solution achieves the most distant and smoothed estimation from the actual simulated sources, while EBB and GS present better results on increasing order. However, the noticeable result here is that the kernel-based *WM-MK* covariance estimator increases the achieved spatial accuracy in all cases. Additionally, all methods present a better power balance among sources.

Analyzing the reconstruction on the temporal domain, using the *COV* approach all three solutions (LOR, EBB and GS) reach similar results for ϵ_t over the examined SNR, and the incorporation of kernel-based covariance estimators enhances the reconstruction in all cases, i.e., the temporal correlation with the time courses are more reliable with the Kernel approaces. Similar to the results obseved for dataset **SDB1**.

As the obtained results are too similar for visual inspection, we replicate the paired t-test in Tab. 4.3 explained above for d_s, in order to assess the statistical significance of ϵ_t results for GS (as it was the one with higher performance). However, in this case the alternative hypothesis states that one of the compared methods mean ϵ_t value is significantly higher. The confidence interval of the ϵ_t estimates is likewise computed as previously, but the alternative hypothesis states that one mean value of ϵ_t in each pair-wise comparison is significantly higher. Adding Gaussian noise (see top row), obtained

values of interval confidence show that *WM-MK* achieves significant differences against other compared methods trough the SNR range. Yet, this advantage decreases in all methods involving kernel-based covariance estimation when considering non-Gaussian noise.

TABLE 4.3: Paired t-test comparing the tested covariance estimations with *GS* with the temporal accuracy index ϵ_t. In the t-test alternative hypothesis, the method named in the right is assumed to have a significantly greater ϵ_t mean value.

SNR		−5			0			5	
	h	t	p	h	t	p	h	t	p
COV vs.IM-SK	1	−10.0034	0 ∗∗	1	−30.2897	0 ∗∗	0	−1.9937	0.0518
COV vs. MED-SK	1	−15.4569	0 ∗∗	1	−16.2527	0 ∗∗	1	5.9091	0 ∗∗
COV vs. WM-MK	1	−22.3932	0 ∗∗	1	−38.5848	0 ∗∗	1	−20.2308	0 ∗∗
IM-SK vs. MED-SK	1	−4.7635	0 ∗∗	1	3.9734	0.0002 ∗∗	1	7.0437	0 ∗∗
IM-SK vs WM-MK	1	−11.4166	0 ∗∗	1	−13.2751	0 ∗∗	1	−16.8647	0 ∗∗
MED-SK vs. WM-MK	1	−2.5426	0.0142∗	1	−14.4398	0 ∗∗	1	−30.6395	0 ∗∗
COV vs.IM-SK	0	0.8483	0.4004	0	−0.7420	0.4617	1	−3.8453	0.0003 ∗∗
COV vs.MED-SK	0	−0.8382	0.4060	1	−2.3881	0.0208∗	1	−2.8423	0.0065 ∗∗
Cov vs. WM-MK	0	−1.9807	0.0533	1	−4.3176	0.0001 ∗∗	1	−5.7455	0 ∗∗
IM-SK vs. MED-SK	0	−1.5620	0.1247	0	−1.4914	0.1423	0	1.8038	0.0774
IM-SK vs WM-MK	1	−2.7328	0.0087 ∗∗	1	−3.7452	0.0005 ∗∗	1	−2.8924	0.0057 ∗∗
MED-SK vs. WM-MK	0	−0.7984	0.4285	0	−1.6892	0.0975	1	−5.5269	0 ∗∗

Overall, the obtained results for **SDB2** coincide with the **SDB1** benchmark and validate our hypothesis that, despite the used inversion method, a better-explained data covariance improves the source localization accuracy.

4.4.3 Performed *RDB1* accuracy

In this case, we use the face perception dataset setting the famous faces stimulus as our target. The source reconstruction obtained by each brain imaging method is averaged from 140 to 180 ms after stimulus to represent the activity corresponding to the N170 peak. Accordingly, a bilateral activation of the fusiform face areas (FFA) is expected, having dominance in the right hemisphere. As recommended in [69], for validation of this specific task, our *ground truth* is extracted from a multimodal pipeline, merging MEG and EEG data acquired for each subject (non-averaged). Achieved multimodal source reconstruction for the first subject of the dataset, shown in the top row of Fig. 4.5, exhibit the highest brain activity near to the bilateral fusiform areas (remarked by red squares), confirming our assumption. The remaining rows display the achieved source reconstructions by the considered brain imaging solutions.

Such as in the visual inspection of the simulated data scenario, GS and EBB exceed the LOR source reconstruction. And again, the proposed *WM-MK* approach improves the brain images of the three solvers with two notorious facts: (i) LOR and GS present
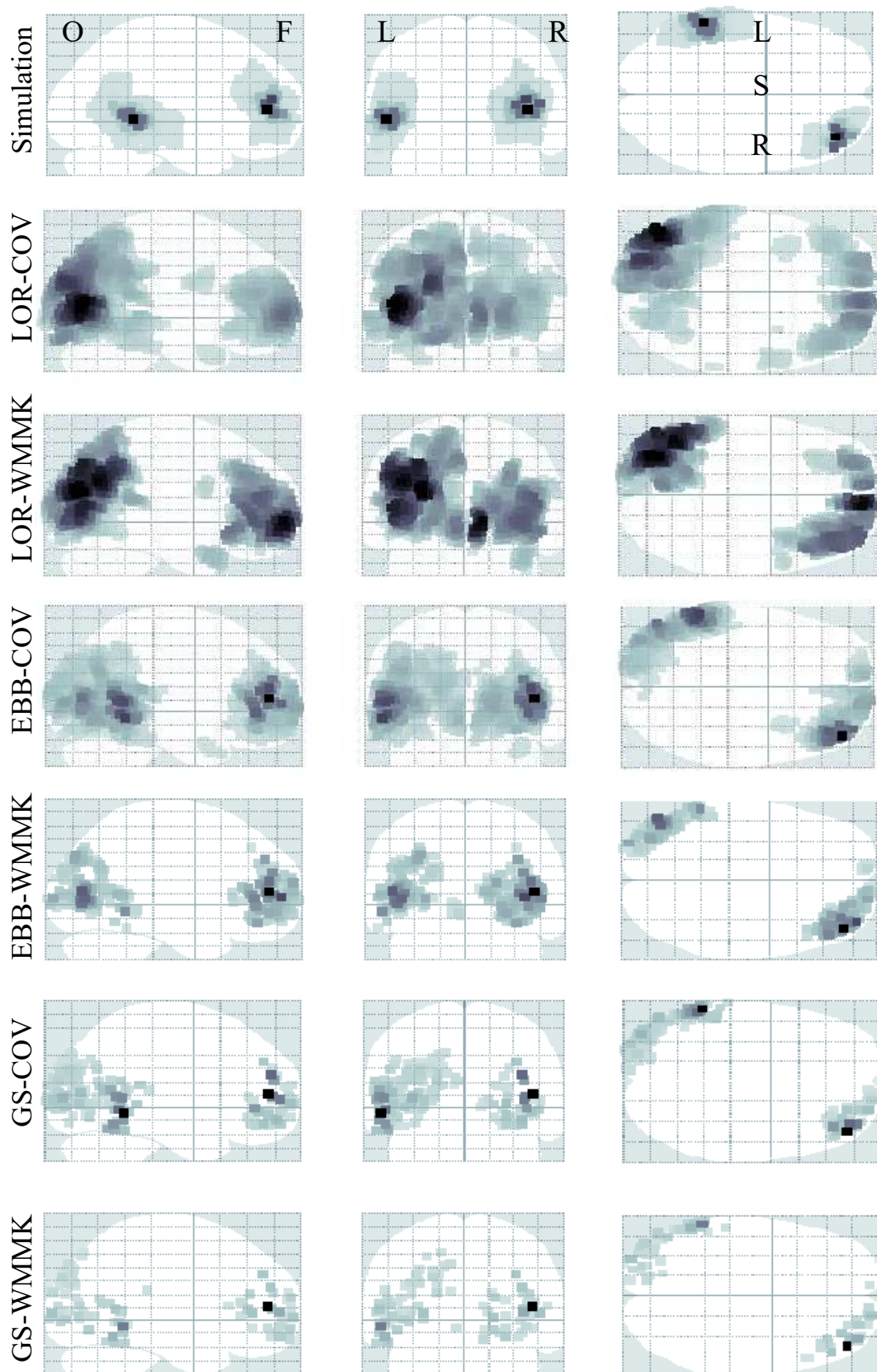
FIGURE 4.4: Glass brain of the simulated neural activity reconstructed by the tested methods at SNR=5 *dB*. Labels LRFO indicates the glass brain views (left, right,frontal , occipital).
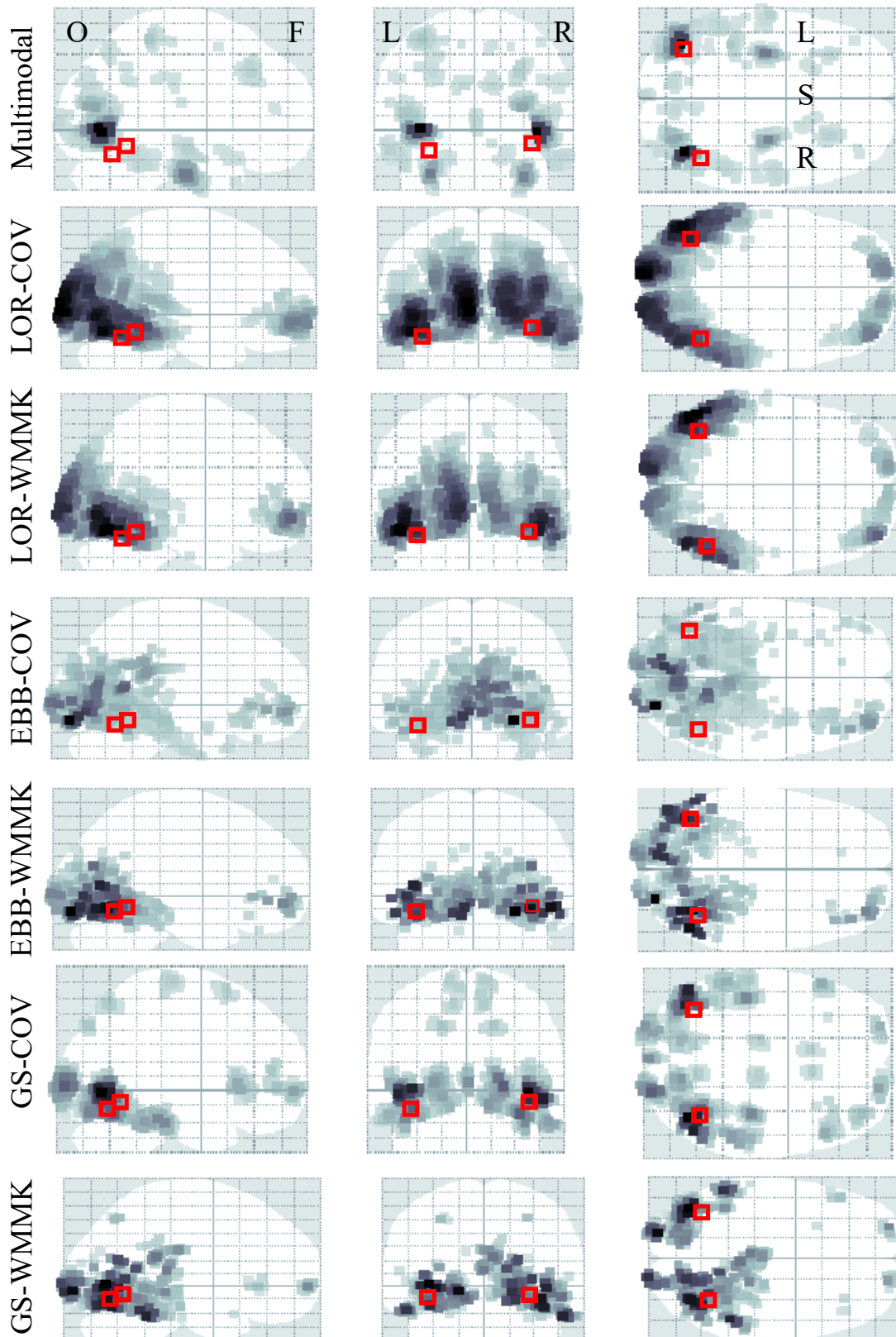
FIGURE 4.5: Glass brains with reconstructed neural activity for the first subject of multimodal faces database. Labels LRFO indicates the glass brain views (Left, right, frontal, occipital). Top row is the multimodal condition taken as ground truth.

lower smoothness with the multi-kernel approach. (ii) The incorporation of the prior kernel information leads to diminishing the spurious activity.

4.4.4 Group Study

One of the main problems of performing group studies with MEG/EEG is the inter subject variability [71], and we hypothesize that most of this variability is caused by the poor accuracy achieved when using the traditional estimator of the covariance of the data (COV). To quantify how the proposed approach can improve results on this aspect, in Fig. 4.6 we matched the *ground-truth* (multimodal pipeline, merging MEG and EEG data acquired for each subject) towards the best performing source reconstruction method (i.e., GS) using the earth-movers distance achieved over each subject on the database (the intuition is that if the multimodal source reconstruction for each subject is close to the source reconstruction with only one modality, the ρ_s should be a low value). This comparison is performed with each subject of the database, although it is not a group comparison, the kernel approach elucidates the hidden activity that a single modality can not see reaching lower values compared with the COV method. Once again, all normalized ρ_s values confirm the findings visualized in Fig. 4.5. That is, by including the *WM-MK* covariance estimator allows outperforming other one-kernel covariance estimates, reaching the lowest ρ_s values, and thus, leading to the best accuracy. Furthermore, the obtained results of the paired t-test, shown in Table 4.4, demonstrate that besides visually enhancing all tested brain imaging solutions, the multi-kernel strategy holds the highest confidence interval of covariance estimation.
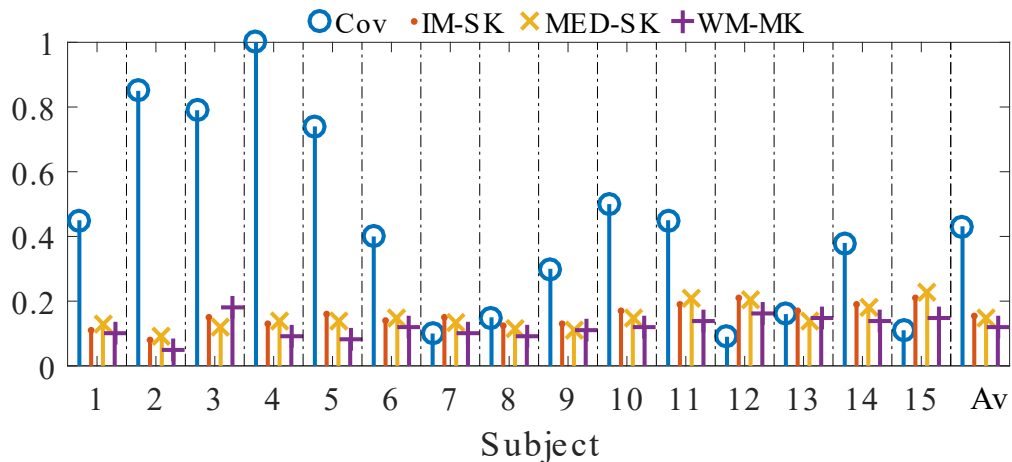
FIGURE 4.6: Earth-movers distance of comparing brain activity reconstructions against our ground-truth. The last column (Av) shows the performance metric averaged across all subject set. Note that the lower values of ρ_s imply that the kernel approaches reveal hidden dynamics in the source space that only one modality cannot observe, i.e., they approximate the multimodal source reconstruction to the single modality results.

TABLE 4.4: *t*-test calculated for earth movers distance

Scenario	ρ_s		
	h	t	p
Cov–IM-SK	1	3.3974	0.0043 ∗∗
Cov-MED-SK	1	3.4721	0.0037 ∗∗
Cov-WM-MK	1	3.8929	0.0016 ∗∗
IM-SK–MED-SK	0	1.0347	0.3184
IM-SK–WM-MK	1	5.3365	0.0001 ∗∗
MED-SK–WM-MK	1	3.4802	0.0037 ∗∗

Finally, we complement this group analysis by assessing the covariance estimator that enables differentiating the most the brain response homogeneity across subjects, applying the kernel alignment metric of Fig. 4.3.4. Note that higher κ, provide better pairwise alignment, and thus, more homogeneous brain responses across subjects (this aligment estimate the similarity across the source reconstruction in all the subjects of the dataset). Fig. 4.7 shows the mean and standard deviation of κ, computed in either space: sensors (see top plot) and sources (bottom plot). As a result, *WM-MK* is the best-considered covariance estimator, because it facilitates distinguishing the paradigm of famous faces as a target stimulus, showing the highest subject responses (i.e., group responses are more related to each other), and reducing the dispersion metric between subjects. Table 4.5 shows the corresponding *t*-test results computed for the κ metrics. The confidence interval of these estimates holds true across subjects, i.e. with the κ measure we are able to judge that the brain responses for this stimulus is in general homogeneous only with the *WM-MK* approach.

TABLE 4.5: Results of *t*-test computed for the κ metrics in sensor and source spaces

Scenario	κ_s Sensor			κ_s Source		
	h	t	p	h	t	p
Cov–IM-SK	1	−20.2945	0 ∗∗	1	−4.0188	0.0011 ∗∗
Cov-MED-SK	1	21.6487	0 ∗∗	1	−9.7663	0 ∗∗
Cov-WM-MK	1	−35.2829	0 ∗∗	1	−10.2391	0 ∗∗
IM-SK–MED-SK	1	−2.1140	0.0356∗	1	−2.7866	0.0138∗
IM-SK–WM-MK	1	−18.3321	0 ∗∗	1	−3.7867	0.0018 ∗∗
MED-SK–WM-MK	1	14.0533	0 ∗∗	1	−4.7868	0.0002∗
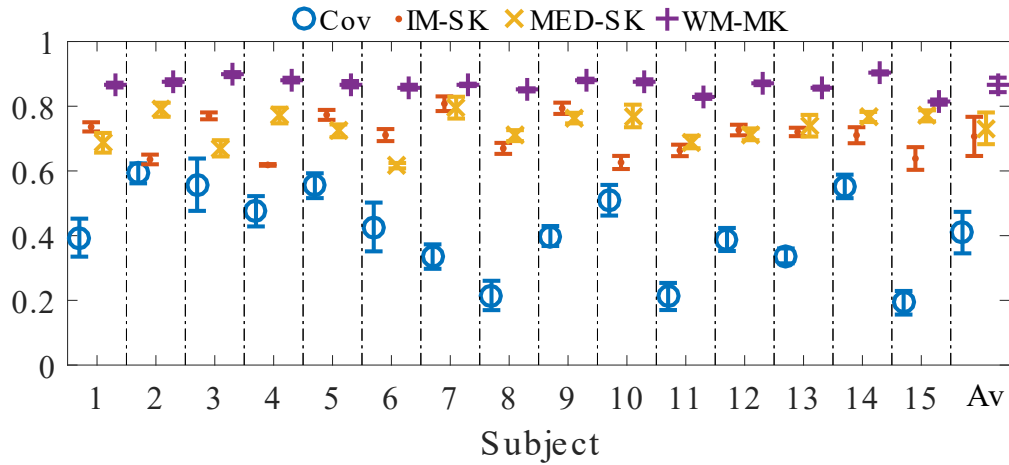
(a) Kernel alignment in sensors
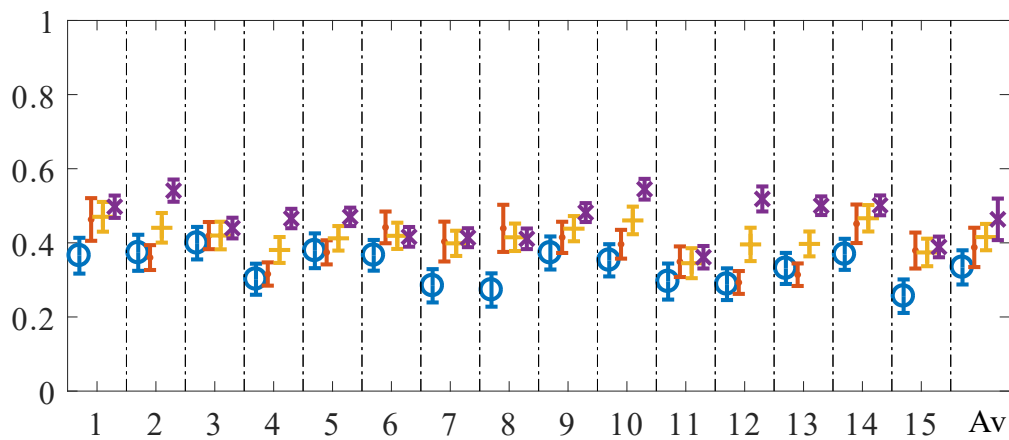
(b) Kernel alignment in sources

FIGURE 4.7: The kernel alignment metric compares the response of subjects to the visual stimulus in sensor and source spaces. Note that the kernel alignment is higher for the kernel prior compared to the covariance prior. The values are as high as twice in some subjects, i.e., the kernel representation is consistent with the group. The last column (Av) shows the performance metrics averaged across all subject set.

4.5 Summary and Discussion

This chapter introduces the kernel-based estimation of MEG/EEG data covariance matrix for improving the neural source reconstruction accuracy. After analysis and validation of several considered approaches to enhanced covariance computation, the following aspects are worth mentioning:

Computation of kernel weights: The implementation of multi-kernel based estimators allows exploiting non-stationary structures with high temporal and spatial complexity. Here, for assessing the influence of enhanced covariance estimation on the tested inversion methods, we compare two borderline situations: *i*) a single kernel function

as the commonly employed method to assess the marginal effect of variable sets (that is, just one weight), and *ii*) Multi-kernel analysis that allows exploiting more complex interactions when having many additive kernel weights.

In the temporal domain the kernel matrix, ruled by its bandwidth, acts as a noise filter highlighting each of the MEG/EEG dynamics and rejecting noise components. As a result, the multi-kernel approach of enhanced covariance estimation achieves the best temporal performance even with non-stationary brain activity under non-Gaussian noise conditions. On spatial domain, the use of kernel-based covariance estimation enables encoding the data relationships elicited by brain activity, appropriately dealing with noise.

Another critical aspect of computing the kernel weights is the adopted cost function. In particular, the use of *KL* divergence provides a straightforward mathematical formulation when employing Gaussian kernels. Moreover, the conic constraints imposed on the optimization problem allow associating each kernel weight with its relevance, providing more interpretable results. Further, optimization is achieved employing gradient descendant strategies, which are easy to implement with a relatively low computational burden.

Estimation of kernel bandwidths: The bandwidth that rules the width of Gaussian shape is a pivotal point of research for kernel filters. Its enhanced estimation results in an RKHS representation that highlights the leading data dynamics. Therefore, several strategies have been proposed for setting this parameter, mostly in accordance with the task at hand. For single kernel functions, we contrast two widely used methods of computation: i) based on information metrics (*IM-SK*), and ii) based on the median of data distances (*MED-SK*). The former one is assumed to promote the data separability, yielding low σ values, though very small σ values make the kernel-based estimation neglect most of the actual data relationships, and the second one with higher σ values should include more noise-induced data relationships. However, both approaches performed similarly in both analyzed domains, meaning that either method is practical.

In the case of a mixture of Gaussian kernels, therefore, we propose to compute the needed parameter set within an extended span, holding the bandwidth setting to enable an adequate multi-kernel representation. The span length depends on a seed value empirically fixed to σ_0 calculated by either approach described above (see Tab. 4.1). The obtained results confirm that the mixed strategy improves the performance of tested brain imaging methods in both temporal and spatial domains. Particularly, Figs. 4.5 and 4.4 demonstrate that *WM-MK* reduces the blurred activity surrounding the active sources, while Fig. 4.7 shows that the estimated time-courses are closer to the actual

ones, even when dealing with non-Gaussian noise. Moreover, *WM-MK* significantly improves the performance of tested brain imaging methods with real data, as shown in Fig. 4.6. Consequently, the mesh of Gaussian kernels is an innovative approach that relaxes the difficult task of estimating a proper kernel bandwidth and improves the performance of the task at hand.

Accuracy of inverse problem solutions: Fig. 4.1 shows that LOR improves the most in both analyzed domains. For instance, in space LOR enhances about 12 points against the linear covariance when using a single kernel, and about 17 points when using the multi-kernel approach. Similar results are observed in time. This behavior is explained as the kernel reduces the LOR noisy and blurred source reconstruction (see Fig. 4.4). Moreover, the accuracy of EBB and GS also rises when including non-linear covariance estimation. Thus, either approach improves about two points with a single kernel and about four points with multiple kernels.

Chapter 5

OPM sensor array localisation using brain imaging based kernel methods

5.1 Introduction

A methodology for reconstructing the true geometry of OPM sensor arrays was presented in Chapter 3. Furthermore, a multi-kernel approach that takes into account the non-stationary nature of MEG/EEG data for brain imaging was presented in Chapter 4. In this chapter, we merge both strategies to create a robust methodology to improve OPM sensor placement.

Additionally, the methodology for reconstructing the geometry of the array had two parameters fixed, the spatial sampling and the head model, and these parameters should be accounted for. The OPM array used in this thesis has 13 sensors to sample the somatosensory cortex (Fig. 3.2), this fixation is set due to the limitation of the number of OPM sensors. However, current arrays have grown up to 32 sensors [72, 73] (as shown in Chapter 3 with the MEG system, our methodology can be implemented with more sensors). The question that remains is, how many sensors are needed for a proper spatial sampling in OPM arrays? Furthermore, we used the single shell model [46] to compute the propagation model. As this parameter may be critical, in this chapter we analyse the effect of using different head models.

Therefore, in this chapter three aspects of the methodology are discussed: the spatial sampling of MEG/EEG and OPM arrays, the justification of the choice of the head

model, and the Kernel-based OPM sensor placement. The methodology uses a multi-kernel approach to estimate the neural source distribution on the cortical surface. the main improvement of including non-stationary information is that the Free Energy becomes convex; therefore, we changed the Metropolis search for a gradient descent optimisation algorithm, allowing us to find the true geometry of the array with approximately the same error, but in much fewer iterations.

5.2 Methodology

In this chapter we merged the methodology to recover the geometry of the OPM array with the multi-kernel approach. The full proposed methodology comprises the following steps: *i)* OPM data collection. *ii)* Variation of the array geometry (the algorithm does not know the true geometry). *iii)* Forward model computation. In this step, we score among three different head models: single shell, MEG local spheres, and single sphere. *iv)* Source estimation (with traditional EBB and with the multi-kernel approach). *v)* Gradient descent with Free Energy as a cost function (repetition of steps *iii* to *v* until convergence). Fig. 5.1 presents a schematic of the methodology used to recover the geometry of the array.
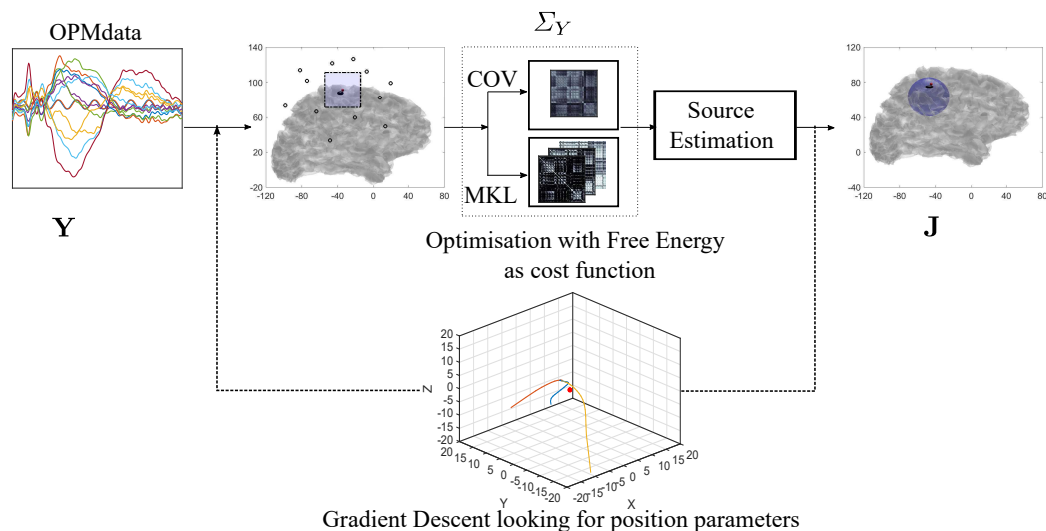
FIGURE 5.1: Methodology to recover the geometry of the OPM array with the MKL approach. The MKL turns the Free Energy in a monotonic function of the location. Thus, the gradient descent could be used to optimise the localisation of the array.

5.2.1 OPMs spatial sampling

The OPM array was designed to sample the somatosensory cortex. In Chapter 3 we show that it is possible to perform a source reconstruction with the OPM array shown in Fig.3.2. Here, we want to probe the reasons for the improvement in orientation (Fig. 3.5(b)) and displacement (Fig. 3.6(b)) precision when using SQUIDs over OPMs. It is possible that although the OPMs observed a larger field change, the spatial sampling of the array on the scalp surface was sub-optimal. We therefore repeated the simulations of Fig.3.5(a) (orientation perturbation error with ideal sensors), but with different potential OPM sensor spacings (and the same number of sensors). Fig. 5.2(a) shows this manipulation. For each sensor array we computed the curves shown in Fig. 3.5(a) (with no gain error added) and calculated the precision or the width of the orientation error curve before intersection with $\Delta F = -3$ (at which point sensor models with greater orientation errors were 20 times less likely than the best model). We did this for putative 4 sensor arrays sharing the same central sensor but rotated around this central axis by 0, 10, 20 and 30 degrees in order to factor out any sampling issues due to the hexagonal arrangement. Fig. 5.2(b) shows that the orientation precision was indeed a function of sensor spacing with optimal spacing (for this simulated source) close to the scanner-cast spacing. Fig. 5.2(c) shows the observed field patterns for the different sensor spacings; it is notable that the closest spacing (1.7 cm) samples the largest field extrema yet is sub-optimal in terms of orientation precision (Fig. 5.2(b)). At the scalp surface the field is much less spatially diffuse requiring much higher density of scalp detectors to effectively achieve the same spatial precision.

In order to maintain the sampling of the standard CTF system (and assuming that on average the CTF sensors are 3 cm further from the surface of an 8 cm radius spherical head); the OPM sensors should be a maximum of 16 mm apart whereas in our array the spacing was ∼30 mm. However, Fig. 5.2(a) show that for this particular region of cortex the scanner-cast sampling was close to optimal; with denser sensor spacing leading to larger magnitude in the sampled field (Fig. 5.2(c)) but less precise orientation estimates (Fig. 5.2(b)). That said, in future oversampling is certainly a safer option than undersampling (and will give rise to a sensitivity increase if nothing else) and although there will be a concomitant increase in cross talk (around 3% for current spacing [4]), this is deterministic and straightforward to model.

(a) Spatial sampling

(b) Sensor spacing
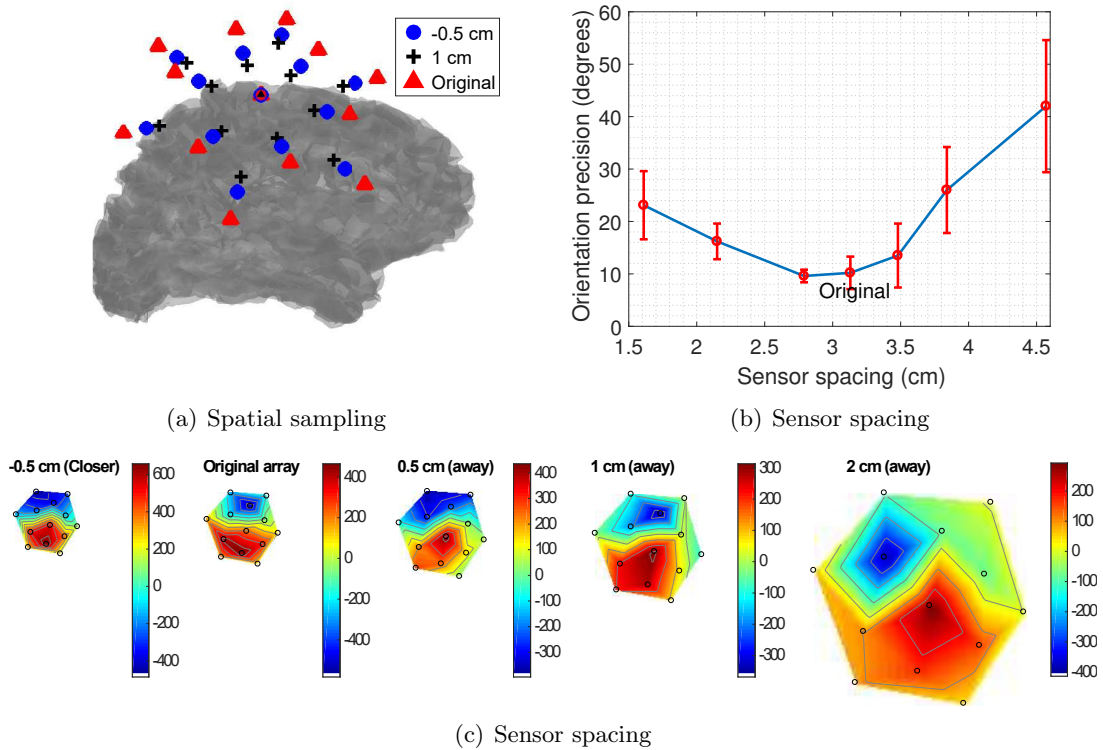
(c) Sensor spacing

FIGURE 5.2: Orientation precision for different array spacings. 5.2(a) The original array (blue circles) with mean sensor separation of approx. 3.1 cm; a denser array with sensors 0.5 cm closer (red triangles) and a coarser array with sensors 1 cm further apart (diamonds, i.e. mean separation approx. 3.6 cm). 5.2(b) Orientation precision (width of perturbation curve at $F = -3$ for different sensor spacings (averaged over concentric rotations of the array of $\{0, 10, 20, 30\}$ degrees). 5.2(c) Field-maps of the simulated source on the different arrays.

5.2.2 Analysing the effect of the head model

In order to justify the choice of the head model used on previous chapters, we used different volume conductor models to explain the single channel OPM data as geometrical distortion. Here, the data remain constant allowing us to directly compare models using Free Energy; in Chapter 3 we examined changes of relative Free Energy (with different sensors and data). In this chapter, we also show these experiments with head models comparison, and also comparing traditional brain imaging methods and kernel functions. As an example, we performed source reconstruction with the single channel OPM data, and used three head models, the single sphere [59], MEG local spheres [74], and single shell model [46], to compare them with relative Free Energy. Additionally, we compared the traditional EBB source reconstruction (Fig. 5.3(a)) with the multi-kernel approach (Fig. 5.3(b)).

As we compared with relative Free Energy, we set the lowest scored model to zero.

The Single sphere model is the one with the lowest Free Energy for both simulations (with traditional EBB and with the multi-kernel approach), followed by the MEG local-spheres and the single Shell model. While both methods (EBB and EBB multi-kernel) present the highest Free Energy with the single shell model, the EBB multi-kernel also surpass the $\Delta F = 3$, where the models are 20 times more likely. This implies that the multi-kernel approach allows a better distinction among the evaluated head models.

(a) EBB covariance   (b) EBB multi-kernel
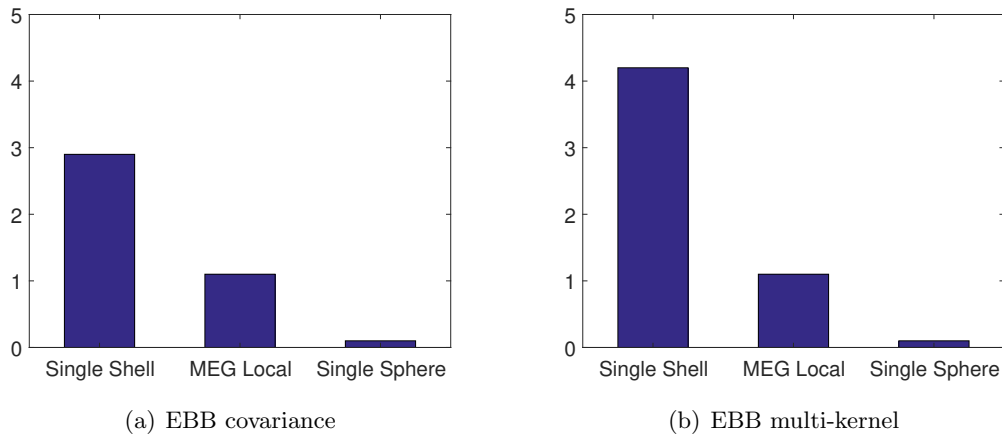
FIGURE 5.3: Influence of the head model. We confirm that among the evaluated head models, the single shell is the most likely. Additionally, the multi-kernel approach presented a better differentiation among models, indicating that the single shell is at least 20 times more likely than the next model.

5.3 Results

5.3.1 Adding sensor orientation error

We added orientation (Fig. 5.4) and position (Fig. 5.5) error to the array with Single Spheres [59], MEG local Spheres [74], and Single Shell [46] models. Fig. 5.4 shows how Free Energy varies as a function of added geometrical noise under different volume conductors for the single channel OPM data. Figures 5.4(a) and 5.4(b) show absolute Free Energy (rather than relative or normalized to maximum). Both panels show that the peak Free Energy (or the most likely model given the data) are sensors without zero orientation error (although the algorithm has no knowledge of true orientation) and that the most likely head model (the one with the highest Free Energy) is the single-shell. These results are in accordance with [75], in that the single shell model outperforms the spherical ones. That said, we were surprised to see such a clear distinction with such a relatively small number (13) of sensors.

Comparing between both strategies (traditional EBB vs multi-kernel approach), the EBB multikernel has the advantage of being more sensitive to changes in orientation error, i.e, the model degrades more rapidly. Also, there is a clearer distinction between the single shell model and its counter part with the multi-kernel approach. In this case, we look at how much the sensor geometry could be degraded before the evidence for the data degrades significantly. The better model (in this case single shell) degrades more rapidly in the presence of geometrical error. The multikernel approach is more sensitive with $7°$ of error, while the traditional EBB is less sensitive with $9°$. This implies that it is easier to distinguish across different geometries with the multi-kernel approach.

(a) EBB covariance

(b) EBB multikernel
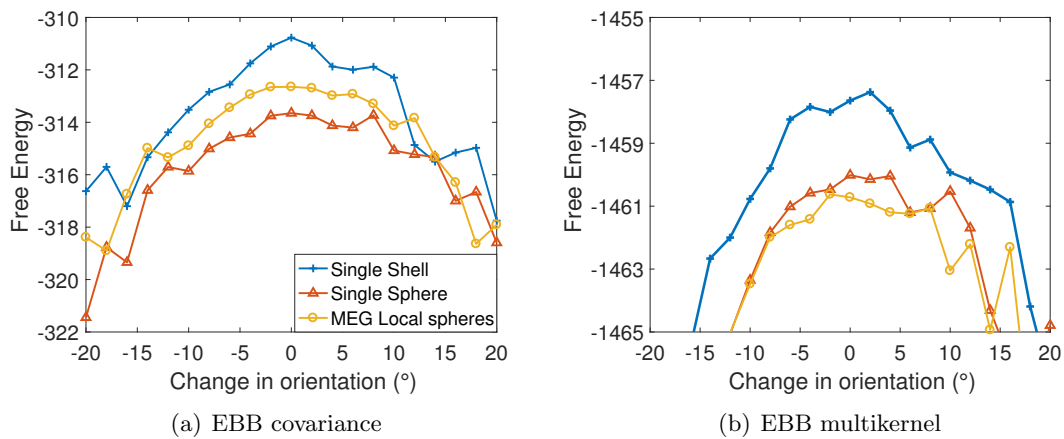
FIGURE 5.4: Adding sensor orientation error

5.3.2 Adding sensor position error

Fig. 5.5 show the same effects for displacements of the sensor array. The single shell model is consistently the most likely (with both strategies, traditional and multi-kernel), but it is notable that even the simplest volume models are sufficient to estimate the true (based on scanner-cast) array geometry. Comparing both strategies, the multi-kernel approach presents the same sensitivity to position error as the traditional EBB (5 mm). However, with position error, the multi-kernel approach is more stable as the function is monotonic (and therefore simpler to optimise).

(a) EBB covariance (b) EBB multikernel
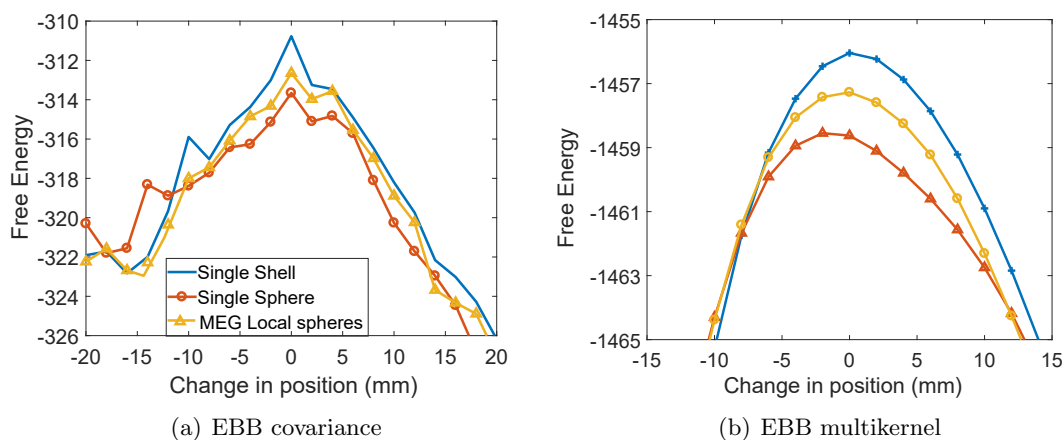
FIGURE 5.5: Adding sensor position error. All methods agree, the most likely head model is the single shell. All the evaluated methods score the true geometry as the most likely, and presented a sensitivity of 5 mm. The multi-kernel approach has the advantage that it provides a monotonic function of the position error.

5.3.3 Model Optimisation

After comparing the traditional EBB and the multi-kernel approach in terms of the sensitivity of the models to orientation and position error, it is established that both strategies perform similarly. Both strategies can recover the true position of the array and the sensitivity to orientation and position error is almost the same (i.e ~ 8 °, and ~ 5 mm).

The practical problem is the same as chapter 3: to demonstrate if it is possible to locate an array with only approximated positional information based only on the field measurements, a volume conductor model, and the cortical geometry. However, we found the Free Energy function measured with the MKL is a monotonic function of the position. Thus, in this chapter we will use the gradient descent optimisation process (as an improvement of the methodology presented in chapter 3)

5.3.4 Gradient Descent

As our findings suggest that the Free Energy function computed with EBB-MKL is a monotonic function of the position, the gradient descent could reduce the number of iterations needed for the optimisation process convergence (compared with the tested Metropolis search). Briefly, the gradient descent used has the Free Energy (F) as a cost function. The goal is:

$$\underset{\mathbf{r}\in\mathbb{R}}{\text{minimise}}\, 1 - F(X) \tag{5.1}$$

where $\mathbf{r} = [x, y, z]$ is the set of position parameters and $F(\mathbf{r}) = F(x, y, z)$ is the Free Energy obtained with the position parameters x, y, z. The following algorithm is used to compute the position parameters:

Data: Prior Positions of $\mathbf{r} = x, y, z$

Result: Final position x, y, z ($\mathbf{r}$)

Initialization;

Choose 5 prior locations (seeds);

Solve the forward problem (compute $\mathbf{L}$) and the inverse problem $\mathbf{J}$ to find $F(x, y, z)$ in
 each location;

Obtain the maximum $F(x, y, z)$ over the priors;

Choose the prior x, y, z with the maximum F;

for $i = 1$ *until convergence* **do**

$\quad x = x - \theta \frac{F(x,y,z) - F(x - \Delta x, y, z)}{\Delta x}$;

$\quad y = y - \theta \frac{F(x,y,z) - F(x, y - \Delta y, z)}{\Delta y}$;

$\quad z = z - \theta \frac{F(x,y,z) - F(x, y, z - \Delta z)}{\Delta z}$;

end

Algorithm 1: Gradient descent algorithm

Here, the numerical derivation is used to compute the gradient. The values of Δx, Δy, Δz are set in 1 mm. The parameter θ is a positive stepsize parameter that is used to ensure that the iteration makes progress towards the solution set of the corresponding problem, here $\theta = 0.3$.

We did the gradient descent procedure using the data from the single-channel OPM array in two ways. Firstly, using a simple 1D search passing over the known location of the sensor array (i.e., both $x, z = 0$ and the optimisation is performed in the y axis). Secondly, by assuming an initial uniform uncertainty over a 64 ($4 \times 4 \times 4$) cm^3 volume a-priori knowledge of sensor array location in any dimension.

5.3.5 Optimisation in one dimension

We used the gradient descent optimiser presented in the previous section. The priors are distributed with $\xi = 80$ mm (i.e. ± 40 mm around the true location, we did this to test the convergence of the algorithm in the convex function of the multi-kernel approach). The gradient descent is performed 40 times for both solvers. Fig. 5.6 shows

the evolution of the position as the number of iterations increases. In each iteration, the position change in the direction of the gradient; in the optimisation with the traditional EBB, only the chains that are in the range $\sim [-8\ 20]$ mm converge (i.e the gradient descent is not an adequate optimiser for the Free Energy computed from the traditional EBB). Comparing the convergence of both algorithms, the EBB-MKL converges for all the chains with a maximum of 13 iterations, even if the priors are set in the extremes (± 40 mm) the algorithm achieves its goal. When both strategies converge, the true geometry error is of 0.34 mm.

(a) Gradient descent with EBB-COV   (b) Gradient descent with EBB-WMMK
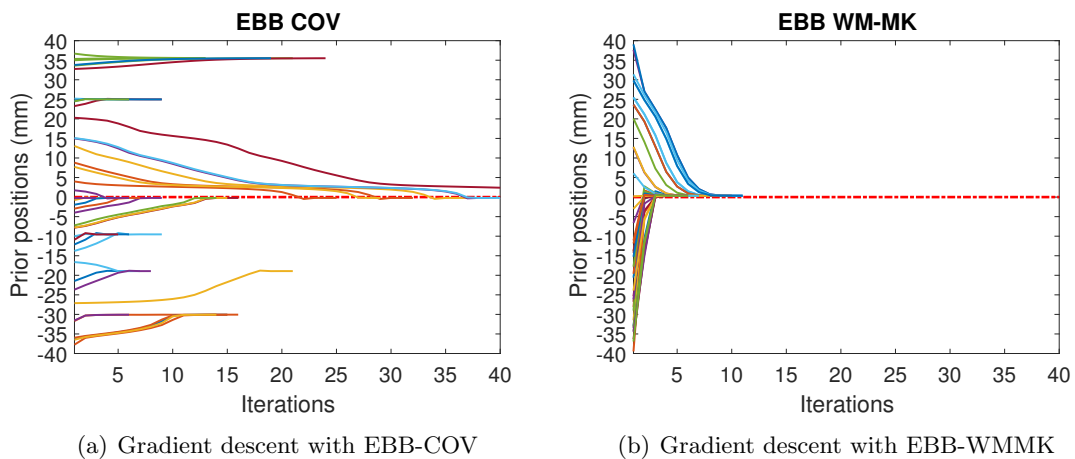
FIGURE 5.6: 1D Optimisation. 5.6(a) shows that the gradient descent is not able to converge with the EBB-COV algorithm in the extended search space $\xi = 80$ mm (i.e. the chains that are outside of the range $[-8\ 20]$ mm got stuck in several local minima). 5.6(b) The gradient descent with the multi-kernel approach reaches the true position of the array with few iterations, it is also capable of finding the true position in a prior space of $\xi = 80$ mm. The average error for the estimation of the array geometry with all the chains for MKL approach is 0.3402 mm

5.3.6   3D Optimisation

Here, we performed the same optimisation with a prior volume of $40 \times 40 \times 40$ mm^3 with the EBB-MKL algorithm that provided a Free Energy convex, ideal to optimise with this procedure. Fig. 5.7 shows four views of the process. The prior cubic volume for the central sensor in the array is presented in Fig. 5.7(a); we performed 4 simulations (for robustness, their starting point is chosen against 5 seeds that are set in the volume for each simulation, the starting point is represented with crosses). The gradient descent process starts with the prior that has the highest Free Energy. As each iteration occurs the array follows the gradient (looking for the global minima of the $1 - F(x, y, z)$). Finally, the algorithm converges in the same point for all the simulations (black point), with an average of 37 iterations (see Appendix B), and the final position is 4.03 mm

displaced from that we expected from the scanner cast (previous computations with traditional EBB and Metropolis search had similar error of ~ 4 mm, with ~ 350 iteraions to converge).

(a) 3D movement

(b) XY view

(c) XZ view

(d) YZ view
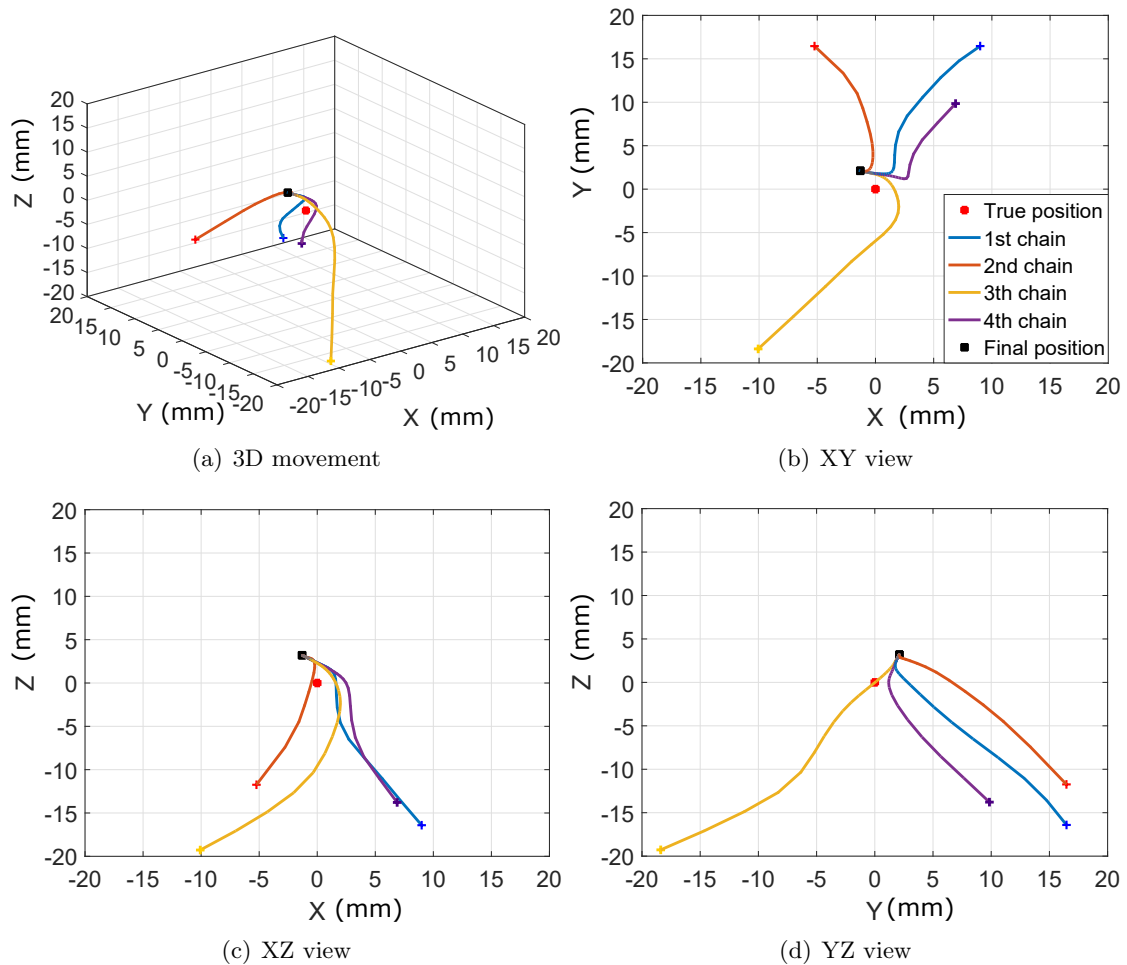
FIGURE 5.7: Four views of the 3D optimisation process with the gradient descent method. Four chains were performed in the 3D optimisation. For each simulation five priors are set in the cube of $40 \times 40 \times 40$ mm^3 and the one with the highest Free Energy is chosen (the starting point in each chain is represented with a cross). All the simulations converged at the same point with 4.04 mm of error (zero is the true position given by scanner-cast, represented here with a red dot, the black dot represents the final position reached for all the chains). The algorithm converges in approximately 40 iterations, compared with the $\sim$350 iterations that Metropolis search uses.

With the gradient descent optimisation process the BMA step is not needed, as the estimation is performed in the point where the optimisation finished. The estimated source has the following MNI location: $[-44.78 \: -47.47 \: 78.12]$, equal to the one computed with the paremeters given by the scanner-cast. With the gradient descent we started with a confident volume of 34.90 cm^3 and at the end we finish with a source with zero localisation error to the source estimate when using the scanner-cast location as ground truth.

5.4 Summary and Discussion

In this chapter the final methodology to find the true geometry of the OPM sensor array was presented. Firstly, the spatial sampling and the head model parameters were discussed. It was found that the tested array was close to the optimal spatial sampling ~ 30 mm. However, a closer array with a mean distance of ~ 2.6 mm could be a safer option. Secondly, we accounted for the effect of the head model in the localisation of the true geometry of the OPM array. We found what was expected, the single shell model performs better than the MEG local spheres and than the single sphere. Moreover, we found that with each head model we can find approximately the true geometry of the array. Solving the inverse problem with the multi-kernel approach, we found that the Free Energy is a monotonic function (easier to optimise). Finally, we tested the methodology to reconstruct the true geometry of the OPM sensor arrays with the EBB multi-kernel approach. With that solver, we were able to use a gradient descent algorithm to optimise the localisation of the array with Free Energy as cost function. With this optimisation, the 1D simulation reached the true geometry with 0.32 mm of error (in chapter 3, the error was 0.44 mm), and the 3D simulation reach the true geometry with 4.03 mm (in chapter 3 the error was ~ 5 mm). Also the methodology with the gradient descent converges in much fewer steps than the Metropolis search. The 1D option with multi-kernel approach converges in 15 iterations, even if the prior is set in $\xi = 80$ mm, while the Metropolis converges in 250 iterations, the reduction is about 94% of the time consumption. For the 3D option, the gradient takes in average 35 iterations to converge, while the Metropolis search takes 370, this implies a time reduction of $\sim 90\%$ of time consumption.

Chapter 6

Conclusions

OPM sensors are rapidly decreasing in size [76] and multi-channel wearable arrays will soon become of clinical use [4, 72, 73]. To date, we have maximized the utility of OPM-MEG data for neural source reconstruction by minimising sensor position uncertainty a priori using scanner-casts (see [5]). In this thesis, we presented a data-driven model optimisation to improve the OPM co-registration with a methodology that recovers the true sensor geometry based on the recorded MEG data and a multi-kernel approach. This methodology reduces the dependence on rigid, time-consuming and somewhat intimidating 3D printed scanner-casts, and potentially gives a way to a more EEG-like system that is flexible, comfortable, and easier to use.

The problem of uncertain sensor placement is not specific to OPM-MEG. In [62] the authors have shown that inaccuracies of EEG electrode coordinates form an error term in the forward model and ultimately in the source reconstruction performance. This error arises from the combination of both intrinsic measurement noise of the digitization device and manual co-registration error when selecting fiducials on anatomical MRI volumes. OPMs pose additional challenges over EEG in that neither orientation nor position will be known in a more flexible set-up. These problems will be yet more acute for the OPMs because the sensitivity to modelling errors is highly dependent on SNR [6, 8, 62].

We derived a framework in which we can compare measurement systems based on their sensitivity to perturbations in sensor geometry. Based on previous SQUID-based studies [14, 16–18, 57], we found that a model of the OPM-MEG data with the true geometry corresponds to higher Free Energy. In this thesis, as we wish to compare between sensor types (and the data were different precluding any direct comparison of Free Energy values), we focused on the sensitivity of the Free Energy to perturbations in the geometry. The rationale is that poor models will be less sensitive to geometrical noise.

6.1 Data-driven model optimisation to improve OPM co-registration

In Chapter 3 an optimisation process necessary to find the OPM sensor array was presented. We found that the Free Energy is a useful metric to judge not only the source reconstruction [19, 60], but also the quality of the forward model. With the Free Energy we were able to score among models, and with a data-driven optimisation find the true geometry of the array.

An important use of this method will be to refine the models of the OPM sensors themselves. For example, the single channel OPM measurements were considerably more sensitive to orientation error than their multi-channel counterparts (Fig. 3.5(b)). We would expect that our models of the multi-channel array improve (by accounting for cross-talk, gain inconsistencies, etc.), so we will observe a tightening of these precision curves. At the moment we can think of two possible reasons why the models of multi-channel data are suboptimal (Fig. 3.5(b)). First, the multichannel system will suffer from cross talk which we estimate to be around 3 % [4]. Second, we made the assumption that the sensor noise covariance matrix $\mathbf{Q}_\Xi$ is a scaled identity matrix (i.e. same noise in all the sensors). For the single (repeated) sensor measurements it is reasonable to assume so, but we estimated the white noise floor (rms $fT/\sqrt{Hz}$) across the multi-channel array varied by around 16%. Third, the OPM measurements took place within the shielded room using an on-board nulling coil to minimize the static field on the sensors; each nulling was followed with a calibration based on the delivery of a pulse of known magnitude through these coils. However, there was no active compensation to minimize the influence of slowly changing magnetic fields during recordings. These fields (of the order of 2 nT peak-to-peak inside the Nottingham University shielded room) give rise to gain changes of the order of 2% [4], which have not been accounted for in this study.

For the simulations, we have assumed that both SQUID and OPM systems had comparable intrinsic noise levels. This was initially a simplification such that system performance could be compared based on geometry alone. We should note that the most recent generation of OPM technology (QuSPin Gen2) has a noise floor or ~7-10fT/rt Hz, comparable to many existing SQUID systems.

With the methodology presented in Chapter 3, we were able to find the geometry of the array with ~ 5 mm of error. The methodology used a Metropolis search to find the true geometry. Furthermore, the Metropolis algorithm is able to converge even if the cost function has local maxima. This is possible because the algorithm sometimes rejects with certain probability the new position parameters that increase the Free Energy.

Allowing such transitions enables the algorithm to escape from local maxima. However, the Metropolis algorithm is slow to converge. The algorithm needs (for the 3D case) ~ 370 iterations. In Chapter 5 we found that the Free energy obtained with the multi-kernel approach is a monotonic and convex function in the prior search space (therefore, easier to optimise).

6.2 Enhanced data covariance estimation:

We introduced a multi-kernel based approach that enhanced the estimation of MEG/EEG data covariance with the purpose of raising the accuracy provided by any brain imaging solver. For encoding different data relationships from measured brain activity, the enhancing method relies on two key contributions: *i*) Input signal decomposition into a mixture of Gaussian functions with weights computed through the *KL* divergence between the kernel and input matrices. To have an adequate multi-kernel representation, we proposed to compute the needed set of kernel bandwidths within an expanded span to encode the wide variability of non-stationary and non-Gaussian input data in temporal and spatial domains. *ii*) The use of *KL* divergence allows associating each kernel weight to its relevance, promoting interpretable results and low computational burden.

The implementation of multi-kernel based estimators allows exploiting nonlinear structures with high temporal and spatial complexity. Here, for assessing the influence of enhanced covariance estimation on the tested inversion methods, we compare two borderline situations: *i)* a single kernel function as the commonly employed method to assess the marginal effect of variable sets, and *ii)* Multi-kernel analysis that allows exploiting more complex interactions when having many additive kernel weights.

In the temporal domain the kernel matrix, ruled by its bandwidth, acts as a noise filter highlighting each of the M/EEG dynamics and rejecting noise components. As a result, the multi-kernel approach of enhanced covariance estimation achieves the best temporal performance even with non-stationary brain activity under non-Gaussian noise conditions. On spatial domain, the use of kernel-based covariance estimation enables encoding the data relationships elicited by brain activity, appropriately dealing with noise.

Validation of *WM-MK* was carried out on non-stationary brain activity (simulated and real-world MEG/EEG data) for evaluating the influence of enhancing the covariance estimation. As a result, the accuracy of the source estimation raises in both analyzed domains, showing that the implementation of multi-kernel based estimators allows exploiting nonlinear structures with high temporal and spatial complexity.

6.3   OPM sensor array localisation using brain imaging based kernel methods

The use of gradient descent optimisation algorithm for localising the OPM sensor array (presented in Chapter 5), improved the methodology proposed in Chapter 3; but this was possible only because the Free Energy obtained with the multi-kernel approach was a monotonical function in the prior search space. Thus, the optimisation was able to converge with 4.02 mm of error, and with $\sim 90\%$ less iterations (compared with the Metropolis search). It must be remarked that it was not possible to guarantee that the positions given by the scanner-cast is 100% accurate. Finally, the source estimate gives the same estimation as the scanner-cast.

One issue, which remains to be tested is how well the estimate of array position will generalise across scalp surface. In may well be that the method is challenged in regions where the forward model is poorly specified (e.g., frontal lobes) or where the generative model is complicated (e.g., the cerebellum)

The experiments performed with OPMs are growing fast. In this thesis we performed a medial nerve stimulation to recover the motor cortex. In [4], the data acquisition was performed while the subject had free movement (over a limited space). The results showed that with the obtained OPM data, it was possible to perform source reconstruction and such modality allowed more movement that EEG (which fails due to muscle artifacts). In [72], the authors performed the first study of human cognition with OPMs. They demonstrated that a wearable OPM system can be used to perform a clinically important language lateralisation paradigm in healthy adults. The results were consistent with fMRI. Importantly, as the OPM array is wearable and motion robust, it could have direct and practical implications in clinical pediatric assessment or in developmental neuroscience. As the system can be used in young children, it can be used to design studies of neurodevelopment language and general cognition or presurgical assessment of 2-8 year old children. Exploiting the movement property, in [77] the authors designed an experiment that involved virtual reality. The authors were able to measure both modulation of alpha-band oscillation by opening and closing the eyes, and the visual evoked field generated by displaying a reversing checkerboard in VR. Moreover, in a VR experiment in which a participant had to look around a wall to view a visual stimulus, the authors showed that MEG signals can be measured and that they map to expected areas of primary visual cortex. Finally, being closer to the head, it is possible to recover deeper structures. In [78], the authors demonstrated that OPMs can be used to study the electrophysiology of the human cerebellum.

6.4 Future work

1. **Integration points:** In this study we have approximated the OPM as a point measurement system. In reality, the volume of the gas exposed to the laser light has maximal dimension of 3mm. This distance is relatively large given that the OPM sensors may now sit ≤ 20 mm from the brain. The addition of appropriate integration points within this volume would be a useful avenue for further study.

2. **Design of new kernels for brain imaging:** We plan to include both pivotal procedures of brain activity analysis (i.e., weight estimation and source reconstruction) into a single Bayesian formulation, attempting to achieve a better coupling between the sensor and source spaces. Besides, the design of kernels, different from Gaussian functions, is to be considered to highlight more complex data relationships.

3. **Use of an improved head model for OPMs:** The increased spatial sampling and sensitivity offered by OPMs will certainly demand more complex head-models. In [6], the authors already showed that small lead-field errors can forsake all potential advantages of OPMs over SQUIDs. Here we have shown that the Nolte single shell model consistently performed better than single and multi- sphere counterparts. As the technology matures, with larger sensor arrays and longer recording times, we will expect to move from the single shell model to more realistically models as the shaped three shell model [75], or the inclusion of more complex models with cerebrospinal fluid, skull spongiosa, and conductivity anisotropy [79].

Appendix A

Optimisation

A.1 Metropolis search

Based on the Metropolis algorithm, we adapted the approach presented in [14] to estimate the true location of the sensor array with respect to the brain:

1. Select a random sample from the prior over possible sensor geometries $h_0 \sim p(h)$ and solve the EBB reconstruction for that geometry. This returns a Free energy value $F(h_0)$.

2. Use a Gaussian distribution to obtain the new array position near to the position computed on the previous step $h_{k-1} : h^{'} \sim \mathcal{N}(h^{'}; h_{k-1}, \sigma^2 I)$.

3. Perform EBB reconstruction on the new location of the array and calculate the ratio r with the new Free energy values:

$$r = \frac{p(\mathbf{Y}|h^{'})ph^{'}}{p(\mathbf{Y}|h_{k-1})p(h_{k-1})} = exp(F(h^{'}) - F(h_{k-1}))\frac{p(h^{'})}{p(h_{k-1})} \tag{A.1}$$

The ratio is given by the comparison of log evidence between the previous reconstruction $p(\mathbf{Y}|h_{k-1})$, and the proposed one $p(Y|h^{'})$, where each is also weighted by the prior. A ratio larger than one means that the proposed geometry $h^{'}$ has more model evidence than the previous one.

4. Take a decision : if $r > 1$ (the new step has higher Free energy), then the new value is higher and accepted $h_k = h^{'}$; if $r < 1$.

5. The new value is compared with a random sample obtained from the uniform distribution: $\beta \sim \mathcal{U}(0,1)$. If $\beta < r$ the parameters are accepted, or rejected

otherwise: $h_k = h_{k-1}$. Allowing such transitions enables the algorithm to escape from local maxima.

6. Return to the second step and repeat until convergence. After an initial burn-in period (first half of data samples), the samples together comprise an approximate posterior distribution over the array locations.

A.2 Bayesian Model Averaging

The following BMA algorithm is used to provide an estimate of the posterior mean $\widehat{J}$. It was set-up for $k = 10,000$ iterations.

1. For the current iteration k, pick a random array geometry from the posterior distribution $h_k \sim p(h|\mathbf{Y})$

2. For the selected array geometry h_k (i.e., its corresponding lead-field matrix), compute the estimated values of the neural activity $\mathbf{J_t}$, and its posterior covariance $\mathbf{\Sigma_k}$.

3. Obtain a normal random variable with mean $\widehat{\mathbf{J}}_\mathbf{k}$ and covariance $(\mathbf{\Sigma_J})_\mathbf{k} : \widehat{\mathbf{J}}_\mathbf{t} \sim \mathcal{N}(\widehat{\mathbf{J}}_\mathbf{t}|\widehat{\mathbf{J}}_\mathbf{k}, (\mathbf{\Sigma_J})_\mathbf{k})$ and save.

4. Update k and go back to step 1. until $k = 10000$.

5. Obtain the mean of the random variables $\widehat{\mathbf{J}} = \Sigma\widehat{\mathbf{J}_\mathbf{k}}/k$

Appendix B

Model optimisation (Metropolis search with MKL)

After comparing the traditional EBB and the multi-kernel approach for computing the sensitivity of the models to orientation and position error, it is established that both strategies perform similarly. Both strategies can recover the true position of the array and the sensitivity to orientation and position error is almost the same (i.e ~ 8 °, and ~ 5 mm).

The practical problem is the same as chapter 3: demonstrate if it is possible to locate an array with only approximate positional information based only on the field measurements, a volume conductor model and the cortical geometry. We do this using the data from the single-channel OPM array in two ways. Firstly, using a simple 1D search passing over the known location of the sensor array. Secondly, by assuming an initial uniform uncertainty over a 64 ($4 \times 4 \times 4$) cm³ volume a-priori knowledge of sensor array location in any dimension. The main difference here, is that the multi-kernel approach is used as the solver to recover the true geometry of the sensor array.

B.0.1   Optimisation in one dimension

We used the Metropolis search algorithm detailed in the appendix. The priors are distributed with $\sigma = 40$ mm (i.e. 20 mm around the true location). The Metropolis search is performed with 600 iterations per chain in four chains. Fig. B.1(a) shows the change in the position error as the number of iterations increases. With each iteration, the distance error diminishes, until reaching a $\sim$0 mm error with 300 iterations. Compared with the traditional approach shown in chapter 3, the convergence delayed $\sim$50 iterations. However, similar to the traditional approach, the algorithm converges. The initial value is

represented with a green point. Through each iteration of the Metropolis search (black points) the position changes (via model comparison Fig. B.1(b)) until convergence (blue point). The error drops with each iteration (Fig. B.1(a)) and after 300 iterations the algorithm oscillates near to the true position. After convergence, the estimated position was 0.38 mm (the traditional EBB was 0.44 mm). The uncertainty (95 percentile) on this geometry estimate is also less than 1 mm. Fig B.1(b) shows, how in each iteration the Free Energy increases until convergence. Fig B.1(c) shows the path followed by the algorithm. Gray dots shows how the algorithm position follows the path of highest Free Energy until reaching an approximate true geometry that is in error of less than 1 mm.

(a) Error

(b) Free Energy

(c) Path

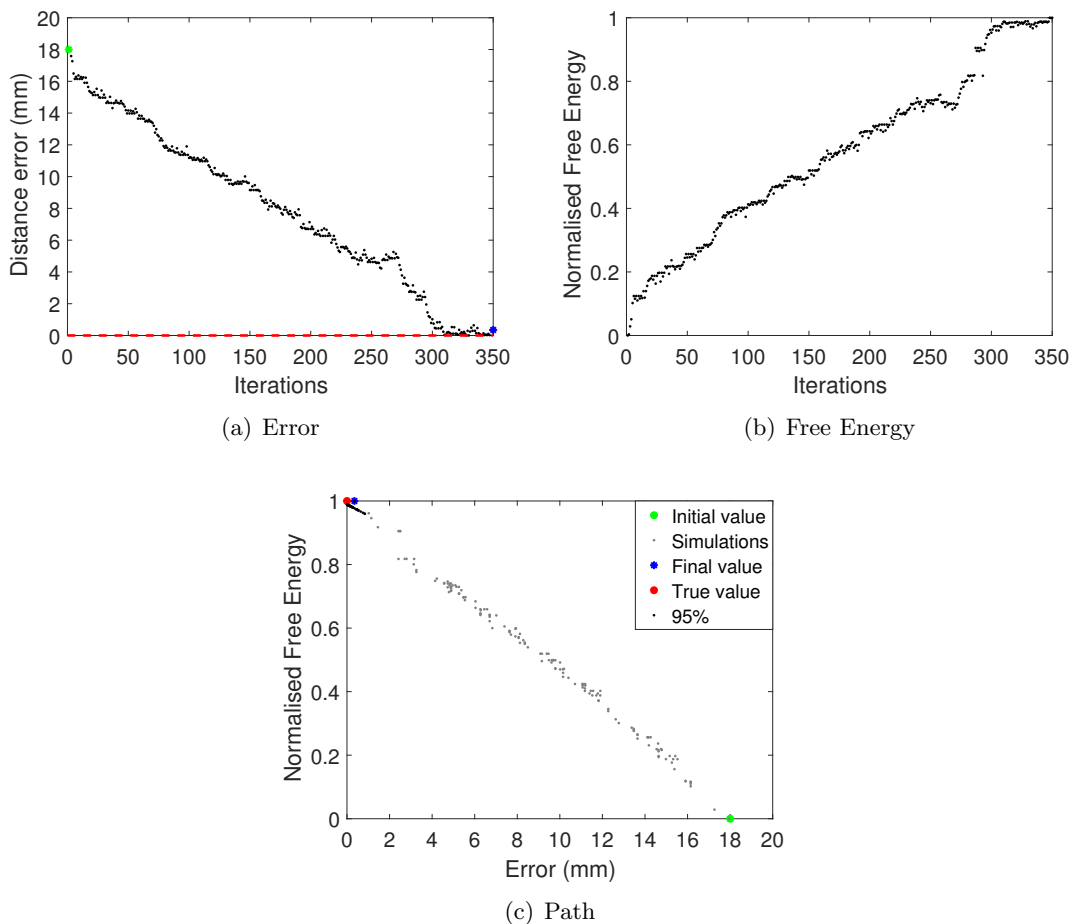FIGURE B.1: 2D Optimisation. B.1(a) shows that with each iteration of the Metropolis algorithm the error diminishes, while in B.1(a) the Free Energy increases with each iteration until convergence. B.1(c) shows the movement of the sensor array across an arch in the XY axis. Initially the error is about 18 mm and finally the algorithm converges. The estimation of the geometry of the array is in error of 0.38 mm.

B.0.2 3D Optimisation

As we indicated in chapter 3, the optimisation in one dimension provides an illustration of the Metropolis process, but it is not practically useful as positional uncertainty is not constrained to lie in one dimension. Here, we performed the same optimisation with a prior volume of $4 \times 4 \times 4$ cm^3. Fig. B.2 shows four views of the process. The prior cubic volume for the central sensor in the array is represented with triangles; 30 priors are set in the volume, and the Metropolis process starts with the one with the highest Free Energy (green point). As each iteration occurs (gray points) the array follows the path with higher Free Energy. Finally, the algorithm converges and the final position is 3.7 mm displaced from that we expected from the scanner cast (previous computations with traditional EBB was in error of 4 mm). The posterior confidence volume on this location was 0.081 cm^3 i.e. near to 600-fold reduction on prior volume.

(a) XY view

(b) XZ view

(c) YZ view
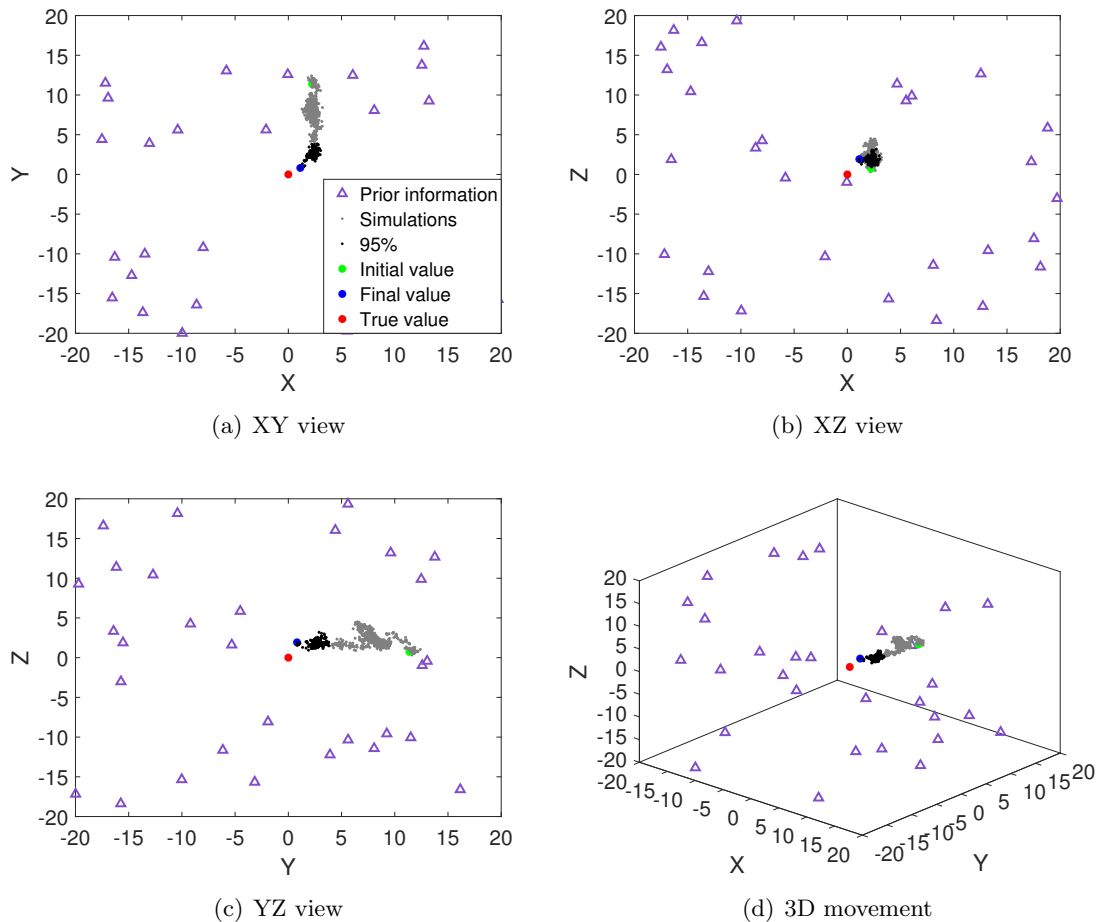
(d) 3D movement

FIGURE B.2: Four views of the 3D optimisation process. In this simulation the largest movement is in the Y axis. First, the priors are set in a cubic distribution of 40 mm with the centre of the array being the true position. One of the priors is chosen, and then the Metropolis process starts. With each iteration the array is moved until convergence. Here, the final position have an error of 3.7 mm.

Estimating the source level activity based on our prior knowledge of sensor position ($4 \times 4 \times 4$ cm^3) gives a distribution of (of peak locations) than can be described by the 95% confidence volume. With the BMA step the model optimisation reduces the confidence volume on peak location from 34.90 cm^3 to 0.051 cm^3. The centre of the optimized confidence volume is 4.2 mm from the source estimate when using the scanner-cast location as ground truth.

Appendix C

Time Analysis

C.1   Single Kernel vs multi-kernel prior

As the computation of the multiple kernel implies more calculations, we performed an analysis of time consumption comparing EBB results with a single kernel for the covariance matrix, with the weighted multiple kernel. One hundred simulations were performed with the covariance matrix and the weighted multiple kernel. Fig. C.1 shows that normalizing the time consumption against the *Cov* method (100 %) the *WM-MK* increases the time consumption on about 30.44 %. However, that is not an issue because in practice they are performed off-line. In terms of time difference, the single kernel last for 5.78 s, and the multiple kernel for 7.54 *s*.
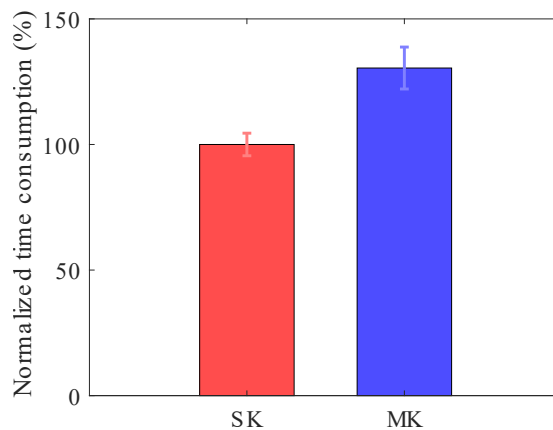

FIGURE C.1: Comparison of time consumption between single kernel and multiple weighted kernels. It increases in average 1.8 s

C.2   Metropolis search vs gradient descent

We found that the Free Energy function computed for the EBB-MKL is a monotonic function, and therefore easier to optimise. In this section, we compare the time cost needed to compute the geometry of the array with the Metropolis search and the gradient descent. Figure C.2 shows the time consumption for the optimisation process in 1D and 3D search for both Metropolis and gradient descent optimisers (for Chapter 3 and chapter 5). We found a reduction of $\sim 90\%$ in the iterations needed for optimise the geometry of the array for the 1D and 3D problems.

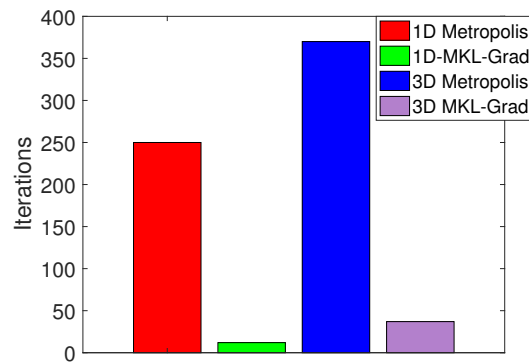FIGURE C.2: Comparison of time consumption for the 1D and 3D optimisation with Metropolis search and gradient descent. While in the 1D optimisation, the Metropolis last for 250 iterations, the gradient descent for the MKL solution takes 12 iterations to converge (i.e. 94.8% fewer iterations). For the 3D optimisation, the Metropolis search last for 370 iterations, while the gradient takes 37, i.e. $\sim 90\%$ fewer iterations.

Bibliography

[1] S. Baillet. Magnetoencephalography for brain electrophysiology and imaging. *Nature Neuroscience*, 20:327–339, 2017.

[2] S. Baillet, J.C. Mosher, and R.M. Leahy. Electromagnetic brain mapping. *IEEE Signal Processing Magazine*, 18(6):14–30, 2001. doi: 10.1109/79.962275.

[3] J.C. Mosher, R.M. Leahy, and P.S. Lewis. Eeg and meg: Forward solutions for inverse methods. *IEEE Transactions on Biomedical Engineering*, 43(3):245–259, 1999.

[4] E. Boto, N. Holmes, J. Leggett, G. Roberts, V. Shah, SS. Meyer, L. Duque-Muoz, KJ. Mullinger, TM. Tierney, S. Bestmann, GR. Barnes, R. Bowtell, and MJ. Brookes. Moving magnetoencephalography towards real-world applications with a wearable system. *Nature*, 555:657–661, 2018.

[5] E. Boto, SS. Meyer, V. Shah, O. Alem, S. Knappe, P. Kruger, TM. Fromhold, M. Lime, PM. Glovera, PG. Morris, R. Bowtell, GR.Barnes, and MJ. Brookes. A new generation of magnetoencephalography: Room temperature measurements using optically-pumped magnetometers. *Neuroimage*, 149:404–414, 2017.

[6] E. Boto, R. Bowtell, P. Krüger, TM. Fromhold, PG. Morris, SS. Meyer, G. Barnes, and M. Brookes. On the potential of a new generation of magnetometers for meg: A beamformer simulation study. *PLoS One*, 11(8), 2016. doi: 10.1371/journal.pone.0157655.

[7] J. Gross, S. Baillet, GR. Barnes, RN. Hensond, A. Hillebrand, O. Jensen, K. Jerbi, V. Litvak, B. Maessh, R. Oostenveld, L. Parkkonen, JR. Taylor, V. Wassenhovek, M. Wibral, and S. Jan-Mathijs Schoffelen. Good practice for conducting and reporting meg research. *Neuroimage*, 65:349–363, 2013.

[8] A. Hillebrand and GR: Barnes. The use of anatomical constraints with meg beamformers. *Neuroimage*, 20(4):2302–2313, 2003.

[9] CJ. Stam. Nonlinear dynamicla analysis of eeg and meg a review of an emerging field. *Clinical Neurophysiology*, 116:2266–2301, 2015.

[10] L. Wang, J. Zhang, L. Zhou, C. Tang, and W. Li. Beyond covariance: Feature representation with nonlinear kernel matrices. *International Conference on Computer Vision*, pages 4570–4578, 2015.

[11] H. Dang, A. Maloof, and M. Romalis. Ultrahigh sensitivity magnetic field and magnetization measurements with an atomic magnetometer. *Applied Physics Letters*, 97(15), 2010. doi: https://doi.org/10.1063/1.3491215.

[12] VK. Shah and RT. Wakai. A compact, high performance atomic magnetometer for biomedical applications. *Phys. Med. Biol.*, pages 8153–8161, 2013. doi: 10.1088/0031-9155/58/22/8153.

[13] C. Pfeiffer, LM. Andersen, D. Lundqvist, M. Hämäläinen, JF. Schneiderman, and R. Oostenveld. Localizing on-scalp meg sensors using an array of magnetic dipole coils. *PLoS ONE*, 13(5), 2018. doi: https://doi.org/10.1371/journal.pone.0191111.

[14] JD. López, WD. Penny, JJ. Espinosa, and GR. Barnes. A general bayesian treatment for meg source reconstruction incorporating lead field uncertainty. *Neuroimage*, 60:1194–1204, 2012.

[15] L. Troebinger, JD. López, A. Lutti, D. Bradbury, S. Bestmann, and GR. Barnes. High precision anatomy for meg. *Neuroimage*, 86:583–901, 2014.

[16] L. Troebinger, JD. López, A. Lutti, S. Bestmann, and GR. Barnes. Discrimination of cortical laminae using meg. *Neuroimage*, 102:885–893, 2014.

[17] Juan D. Martinez-Vargas, Gregor Strobbe, Kristl Vonck, Pieter van Mierlo, and German Castellanos-Dominguez. Improved localization of seizure onset zones using spatiotemporal constraints and time-varying source connectivity. *Frontiers in Neuroscience*, 11:156, 2017. doi: 10.3389/fnins.2017.00156. URL https://www.frontiersin.org/article/10.3389/fnins.2017.00156.

[18] SS. Meyer, H. Rossitera, MJ. Brookes, MW. Woolrich, S. Bestmann, and GR. Barnes. Using generative models to make probabilistic statements about hippocampal engagement in meg. *Neuroimage*, 49:468–482, 2017.

[19] K. Friston, L. Harrison, J. Daunizeau, S. Kiebel, C. Phillips, N. Trujillo-Barreto, R. Henson, G. Flandin, and J. Mattoutf. Multiple sparse priors for the m/eeg inverse problem. *Neuroimage*, 39:1104–1120, 2008.

[20] A. Gelman, JB. Carlin, HS. Stern, DB. Dunson, A. Vehtari, and DB. Rubin. *Bayesian Data Analysis*. Chapman and Hall/CRC, 2000.

[21] JD. Martínez-Vargas, JD López, A. Baker, G. Castellanos-Dominguez, MW. Wool-rich, and GR. Barnes. Non-linear parameter estimates from non-stationary meg data. *Front. Neurosci*, 10(316), 2016. doi: 10.3389/fnins.2016.00366.

[22] G. Pillonettoa, F. Dinuzzo, T. Chen, and Nicolao L. Ljung GD. Kernel methods in system identification, machine learning and function estimation: A survey. *Automatica*, 50:657–682, 2014.

[23] J. Shawe-Taylor and N. Cristianini. *Kernel Methods for Pattern Analysis*. Cambridge University Press, 2004.

[24] T. Gärtner. A survey of kernels for structured data. *ACM SIGKDD Explorations Newsletter*, 5:49–58, 2003.

[25] L. A. Belanche Muñoz. Developments in kernel design. In *2013 proceedings: European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning*, volume 5, pages 369–378, 2013.

[26] W. Liu, JC. Principe, and S. Haykin. *Kernel adaptive filtering: a comprehensive introduction*. John Wiley & Sons, 2011.

[27] X. Zhang, J. Li, and H. Yu. Local density adaptive similarity measurement for spectral clustering. *Pattern Recognition Letters*, 32(2):352–358, 2011.

[28] AM. Álvarez-Meza, JA. Lee, M. Verleysen, and G. Castellanos-Dominguez. Kernel-based dimensionality reduction using renyi's α-entropy measures of similarity. *Neurocomputing*, 222:36–46, 2017.

[29] M. Gonen and E. Alpaydin. Multiple kernel learinig algorithms. *Journal of Machine Learning Research*, 12:2211–2268, 2011.

[30] K. Friston, J. Mattout, N. Trujillo-Barreto, J. Ashburner, and W. Penny. Variational free energy and the laplace approximation. *Neuroimage*, 34:220–234, 2007.

[31] F. Lotte, M. Congedo, A. Lécuyer, F. Lamarche, and B. Arnaldi. A review of classification algorithms for eeg-based braincomputer interfaces. *Journal of Neural Engineering*, 4:R1–R13, 2007. doi: 10.1088/1741-2560/4/2/R01.

[32] A. Schnitzler and J. Gross. Normal and pathological oscillatory communication in the brain. *Nature Reviews Neuroscience*, 6:285–296, 2005.

[33] Am Beres. Time is of the essence: A review of electroencephalography (eeg) and event-related brain potentials (erps) in language research. *Appl Psychophysiol Biofeedback.*, 42:247–255, 2017. doi: 10.1007/s10484-017-9371-3.

[34] S. Noachtar and S. Rémi. The role of eeg in epilepsy: a critical review. *Epilepsy Behavior*, 15:22–33, 2009.

[35] E. Basar and B. Güntekin. A review of brain oscillations in cognitive disorders and the role of neurotransmitters. *Brain research*, 1235:172–193, 2008.

[36] E. Niedermeyer and L. Silva. *Electroencephalography: Basic principles, clinicl applications, and related fields.* Lippincot Williams & Wilkins, 2005.

[37] V Sakkalis. Review of advanced techniques for the estimation of brain connectivity measured with eeg/meg. *Computers in Biology and Medicine*, 41:1110–1117, 2011.

[38] MP. Ledbetter, IM. Savukov, VM. Acosta, D. Budker, and MV. Romalis. Spin-exchange-relaxation-free magnetometry with cs vapor. *Physical Review A*, 77, 2008.

[39] Dupont-Roc, S. Haroche, and C. Cohen-Tannoudji. Detection of very weak magnetic fields (10^{-9} gauss) by ^{8}7rb zero-field level crossing resonances. *Physics Letters A*, 28:638–639, 1969.

[40] W. Happer and H. Tang. Spin-exchange shift and narrowing of magnetic resonance lines in optically pumped alkali vapors. *Phys. Rev. Lett*, 31, 1973.

[41] R. Benumof. Optical pumping theory and experiments. *Am. J. Phys.*, 33:151–165, 1965.

[42] C. Cohen-Tannoudji, J. DuPont-Roc, S. Haroche, and F. Laloë. Detection of the static magnetic field produced by the oriented nuclei of optically pumped ^{3}he gas. *Phys. Rev. Lett*, 22(758), 1969.

[43] A. Kastler. The hanle effect and its use for the measurements of very small magnetic fields. *Nuclear Instruments and Methods*, 110:259–265, 1973.

[44] DB. Geselowitz. On the magnetic field generated outside an inhomogeneous volume conductor by internal current sources. *IEEE Trans. Magn.*, 6:346–347, 1970.

[45] AM. Dale and MI. Sereno. Improved localizadon of cortical activity by combining eeg and meg with mri cortical surface reconstruction: a linear approach. *Journal of cognitive neuroscience*, 5:162–176, 1993.

[46] G. Nolte. The magnetic lead field theorem in the quasi-static approximation and its use for magnetoencephalography forward calculation in realistic volume conductors. *Phys. Med. Biol*, 48:3637–3652, 2003.

[47] R. Grech, T. Cassar, J. Muscat, K. Camilleri, M. Zervakis S. Fabri, P. Xanthopoulos, V. Sakkalis, and B. Vanrumste. Review on solving the inverse problem in EEG source analysis. *Journal of Neuroengineering and Rehabilitation*, 5, 2008.

[48] K. Sekihara, S. Nagarajan, D. Poeppel, and Y. Miyashita. Reconstructing spatio-temporal activities of neural sources using an meg vector beamformer technique. *IEEE Transactions on Biomedical Engineering*, 48:760–3771, 2001.

[49] P. Belardinelli, E. Ortiz, G. Barnes, U. Noppeney, and H. Preissl. Source reconstruction accuracy of MEG and EEG bayesian inversion approaches. *PLoS ONE*, 2:1–16, 2012.

[50] DA. Engemann and A. Gramfort. Automated model selection in covariance estimation and spatial whitening of MEG and EEG signals. *Neuroimage*, 108:328–342, 2015.

[51] D. Wipf and S. Nagarajan. A unified bayesian framework for MEG/EEG source imaging. *Neuroimage*, 44:947–966, 2009.

[52] R. Henson, G. Flandin, K. Friston, and J. Mattout. A parametric empirical bayesian framework for fMRI-constrained MEG/EEG source reconstruction. *Human Brain Mapping*, 31:1512–1531, 2010.

[53] R. Henson, J. Mattout, C. Phillips, and K. Friston. Selecting forward models for meg source-reconstruction using model-evidence. *Neuroimage*, 46:168–176, 2009. doi: 10.1016/j.neuroimage.2009.01.062.

[54] J. López, J. Espinosa V. Litvak, K. Friston, and G. R. Barnes. Algorithmic procedures for bayesian meg/eeg source reconstruction in spm. *Neuroimage*, 84:476–487, 2014.

[55] RN. Henson, DG. Wakeman, V. Litvak, and KJ. Friston. A parametric empirical bayesian framework for the eeg/meg inverse problem: Generative models for multi-subject and multi-modal integration. *Front Hum Neurosci*, 5, 2011. doi: 10.3389/fnhum.2011.00076.

[56] WD. Penny. Comparing dynamic causal models using aic, bic and free energy. *Neuroimage*, 59:319330, 2011. doi: 10.1016/j.neuroimage.2011.07.039.

[57] C. Stevenson, M. Brookes, JD. López, L. Troebinger, J. Mattout, W. Penny, P. Morris, A. Hillebrand, R. Henson, and G. Barnes. Does function fit structure? a ground truth for non-invasive neuroimaging. *Neuroimage*, 94:8995, 2014. doi: 10.1016/j.neuroimage.2014.02.033.

[58] J. Iivanainen, M. Stenroos, and L. Parkkonen. Measuring meg closer to the brain: Performance of on-scalp sensor arrays. *Neuroimage*, 147:542–553, 2017.

[59] MS. Hämäläinen and J. Sarvas. Feasibility of the homogeneous head model in the interpretation of neuromagnetic fields. *Phys. Med. Biol.*, 32:91–97, 1987.

[60] JD. López, F. Valencia, G. Flandin, W. Penny, and GR. Barnes. Reconstructing anatomy from electro-physiological data. *Neuroimage*, 163:480–486, 2017.

[61] H. Wiström, J. Huttunen, A. Korvenoja, J. Virtanen, O. Salonen, H. Aronen, and R.J. Ilmoniemi. Effects of interstimulus interval on somatosensory evoked magnetic fields (sefs): A hypothesis concerning sef generation at the primary sensorimotor cortex. *Neurophysiology*, 100(6):479–487, 1996.

[62] SS. Dalal, S. Rampp, F. Willomitzer, and S. Ettl. Consequences of eeg electrode position error on ultimate beamformer source reconstruction performance. *Front. Neurosc*, 2014. doi: 10.3389/fnins.2014.00042.

[63] C. A. Torres-Valencia, M. C. J. Santamaria, and M. A. Alvarez. Kernel temporal enhancement approach for loreta source reconstruction using eeg data. *Engineering in Medicine and Biology Society (EMBC), 38th Annual International Conference of the IEEE*, 2016.

[64] Y. Ying, K. Huang, and C. Campbell. Enhanced protein fold recognition through a novel data integration approach. *BMC bioinformatics*, 10:267, 2009.

[65] J. Kaipio and E. Somersalo. *Statistical and computational inverse problems*. Springer Science & Business Media, 2006.

[66] S. Haufe and A. Ewald. A simulation framework for benchmarking eeg-based brain connectivity estimation methodologies. *Brain topography*, pages 1–18, 2016.

[67] Y. Huang, L. C. Parra, and S. Haufe. The new york head– a precise standardized volume conductor model for eeg source localization and tES targeting. *Neuroimage*, 140:150–62, 2016.

[68] S. Castano-Candamil, J. Hohne, A. Xing-Wei J. Martinez-Vargas, G. Castellanos-Dominguez, and S. Haufe. Solving the EEG inverse problem based on space-time-frequency structured sparsity constraints. *NeuroImage*, 118:598–612, 2015.

[69] D.G. Wakeman and R. N Henson. A multi-subject, multi-modal human neuroimaging dataset. *Scientific data*, 2, 2015.

[70] A. M. Álvarez-Meza, D. Cárdenas-Peña, and G. Castellanos-Dominguez. Unsupervised kernel function building using maximization of information potential variability. *Iberoamerican Congress on Pattern Recognition, Springer*, pages 335–342, 2014.

[71] V. Litvak and K.J. Friston. Electromagnetic source reconstruction for group studies. *Neuroimage*, 42:1490–1498, 2008.

[72] TM. Tierney, N. Holmes, SS. Meyer, G. Roberts, J. Legget, S. Buck, L. Duque-Muñoz, V. Litvak, S. Bestmann, T. Baldeweg, R. Bowtell, MJ. Brookes, and GR. Barnes. Cognitive neuroscience using wearable magnetometer arrays: Non-invasive assessment of language function. *Neuroimage*, 181:513–520, 2018.

[73] L. Chin-Hsuan, TM. Tierney, N. Holmes, E. Boto, J. Legget, R. Bowtell, MJ. Brookes, RC. Miall, and GR. Barnes. Using optically-pumped magnetometers to measure magnetoencephalographic signals in the human cerebellum. *biorxiv*, 2018. doi: https://doi.org/10.1101/425447.

[74] M.X. Huang, J.C. Mosher, and R. Leahy. A sensor weighted overlapping-sphere head model and exhaustive head model comparison for meg. *Phys. Med. Biol*, 44: 423–440, 1999.

[75] M. Stenros, A. Hunold, and J. Haueisen. Comparison of three-shell and simplified volume conductor models in magnetoencefalography. *Neuroimage*, 94:337–348, 2014.

[76] O. Alem, AM. Benison, DS. Barth, J. Kitching, and S. Knappe. Magnetoencephalography of epilepsy with a microfabricated atomic magnetrode. *J. Neurosci*, 34(43):14324 –7, 2014.

[77] G. Roberts, N. Holmes, N. Alexander, E. Boto, J. Leggett, RM. Hill, V. Shah, M. Rea, R. Vaughan, EA. Maguire, K. Kessler, S. Beebe, M. Fromhold, GR. Barnes, R. Bowtell, and MJ. Brookes. Towards opm-meg in a virtual reality environment. *Neuroimage*, pages 408–417, 2019. doi: 10.1016/j.neuroimage.2019.06.010.

[78] L. ChinHsuan, TM. Tierney, N. Holmes, E. Boto, J. Leggett, S. Bestmann, R. Bowtell, MJ. Brookes, GR. Barnes, and C. Miall. Using optically pumped magnetometers to measure magnetoencephalographic signals in the human cerebellum. 2019. doi: 10.1113/JP277899.

[79] J. Vorwerk, J.H. Cho, S. Rampp, H. Hamer, T.R. Knösche, and C.H. Wolters. A guideline for head volume conductor modeling in eeg and meg. *Neuroimage*, 100: 590–607, 2014.