Mapping, timing and tracking cortical activations with MEG and EEG: Methods and application to human vision

Alexandre Gramfort

PhD THESIS

prepared at
INRIA Sophia Antipolis

and presented at
Graduate School of Telecom ParisTech

A dissertation submitted in partial fulfillment
of the requirements for the degree of

DOCTOR OF SCIENCE
Specialized in Signal and Image Processing

Mapping, timing and tracking cortical activations with MEG and EEG: Methods and application to human vision

Alexandre GRAMFORT

Advisors	Dr. Maureen Clerc	ENPC / INRIA Sophia Antipolis, France
	Pr. Olivier Faugeras	INRIA Sophia Antipolis, France
Reviewers	Pr. Matti Hämäläinen	MGH/MIT/HMS Martinos Center, Boston, USA
	Dr. Rémi Gribonval	IRISA, Rennes, France
Examiners	Dr. Sylvain Baillet	Medical College of Wisconsin, Milwaukee, USA
	Pr. Eric Moulines	Telecom ParisTech, Paris, France
Invited scientist	Dr. Elsa Angelini	Telecom ParisTech, Paris, France

École Doctorale
d'Informatique,
Télécommunications
et Électronique de Paris

THÈSE

présentée pour obtenir le grade de Docteur
de l'École Nationale Supérieure des Télécommunications
Spécialité : Signal et Image

Alexandre GRAMFORT

Localisation et suivi d'activité fonctionnelle cérébrale en électro et magnétoencéphalographie : Méthodes et applications au système visuel humain

Soutenue le 12 Octobre 2009 devant le jury composé de:

Composition du jury:

Directeurs	Dr. Maureen Clerc	ENPC / INRIA Sophia Antipolis, France
	Pr. Olivier Faugeras	INRIA Sophia Antipolis, France
Rapporteurs	Pr. Matti Hämäläinen	MGH/MIT/HMS Martinos Center, Boston, USA
	Dr. Rémi Gribonval	IRISA, Rennes, France
Examinateurs	Dr. Sylvain Baillet	Medical College of Wisconsin, Milwaukee, USA
	Pr. Eric Moulines	Telecom ParisTech, Paris, France
Invités	Dr. Elsa Angelini	Telecom ParisTech, Paris, France

A mon grand-père.

ABSTRACT

The overall aim of this thesis is the development of novel electroencephalography (EEG) and magnetoencephalography (MEG) analysis methods to provide new insights to the functioning of the human brain. MEG and EEG are non-invasive techniques that measure outside of the head the electric potentials and the magnetic fields induced by the neuronal activity, respectively. The objective of these functional brain imaging modalities is to be able to localize in space and time the origin of the signal measured. To do so very challenging mathematical and computational problems needs to be tackled. The first part of this work proceeds from the biological origin the M/EEG signal to the resolution of the forward problem. Starting from Maxwell's equations in their quasi-static formulation and from a physical model of the head, the forward problem predicts the measurements that would be obtained for a given configuration of current generators. With realistic head models the solution is not known analytically and is obtained with numerical solvers. The first contribution of this thesis introduces a solution of this problem using a symmetric boundary element method (BEM) which has an excellent precision compared to alternative standard BEM implementations. Once a forward model is available the next challenge consists in recovering the current generators that have produced the measured signal. This problem is referred to as the inverse problem. Three types of approaches exist for solving this problem: parametric methods, scanning techniques, and image-based methods with distributed source models. This latter technique offers a rigorous formulation of the inverse problem without making strong modeling assumptions. However, it requires to solve a severely ill-posed problem. The resolution of such problems classically requires to impose constraints or priors on the solution. The second part of this thesis presents robust and tractable inverse solvers with a particular interest on efficient convex optimization methods using sparse priors. The third part of this thesis is the most applied contribution. It is a detailed exploration of the problem of retinotopic mapping with MEG measurements, from an experimental protocol design to data exploration, and resolution of the inverse problem using time frequency analysis. The next contribution of this thesis, aims at going one step further from simple source localization by providing an approach to investigate the dynamics of cortical activations. Starting from spatiotemporal source estimates the algorithm proposed provides a way to robustly track the "hot spots" over the cortical mesh in order to provide a clear view of the cortical processing over time. The last contribution of this work addresses the very challenging problem of single-trial data processing. We propose to make use of recent progress in graph-based methods in order to achieve parameter estimation on single-trial data and therefore reduce the estimation bias produced by standard multi-trial data averaging. Both the source code of our algorithms and the experimental data are freely available to reproduce the results presented. The retinotopy project was done in collaboration with the LENA team at the hôpital La Pitié-Salpêtrière (Paris).

Keywords:
Neuroimaging, magnetoencephalography (MEG), electroencephalography (EEG), human vision, retinotopy, boundary element method, inverse problem, convex optimization, sparse regression, single-trial analysis, graph cuts.

RÉSUMÉ

Cette thèse est consacrée à l'étude des signaux mesurés par Electroencéphalographie (EEG) et Magnétoencéphalographie (MEG) afin d'améliorer notre compréhension du cerveau humain. La MEG et l'EEG sont des modalités d'imagerie cérébrale non invasives. Elles permettent de mesurer, hors de la tête, respectivement le potentiel électrique et le champ magnétique induits par l'activité neuronale. Le principal objectif lié à l'exploitation de ces données est la localisation dans l'espace et dans le temps des sources de courant ayant généré les mesures. Pour ce faire, il est nécessaire de résoudre un certain nombre de problèmes mathématiques et informatiques difficiles. La première partie de cette thèse part de la présentation des fondements biologiques à l'origine des données M/EEG, jusqu'à la résolution du problème direct. Le problème direct permet de prédire les mesures générées pour une configuration de sources de courant donnée. La résolution de ce problème à l'aide des équations de Maxwell dans l'approximation quasi-statique passe par la modélisation des générateurs de courants, ainsi que de la géométrie du milieu conducteur, dans notre cas la tête. Cette modélisation aboutit à un problème direct linéaire qui n'admet pas de solution analytique lorsque l'on considère des modèles de tête réalistes. Notre première contribution porte sur l'implémentation d'une résolution numérique à base d'éléments finis surfaciques dont nous montrons l'excellente précision comparativement aux autres implémentations disponibles. Une fois le problème direct calculé, l'étape suivante consiste à estimer les positions et les amplitudes des sources ayant généré les mesures. Il s'agit de résoudre le problème inverse. Pour ce faire, trois méthodes existent: les méthodes paramétriques, les méthodes dites de "scanning", et les méthodes distribuées. Cette dernière approche fournit un cadre rigoureux à la résolution de problème inverse tout en évitant de faire de trop importantes approximations dans la modélisation. Toutefois, elle impose de résoudre un problème fortement sous-contraint qui nécessite de fait d'imposer des *a priori* sur les solutions. Ainsi la deuxième partie de cette thèse est consacrée aux différents types d'*a priori* pouvant être utilisés dans le problème inverse. Leur présentation part des méthodes de résolution mathématiques jusqu'aux détails d'implémentation et à leur utilisation en pratique sur des tailles de problèmes réalistes. Un intérêt particulier est porté aux *a priori* induisant de la parcimonie et conduisant à l'optimisation de problèmes convexes non différentiables pour lesquels sont présentées des méthodes d'optimisation efficaces à base d'itérations proximales. La troisième partie porte sur l'utilisation des méthodes exposées précédemment afin d'estimer des cartes rétinotopiques dans le système visuel à l'aide de données MEG. La présentation porte à la fois sur les aspects expérimentaux liés au protocole d'acquisition jusqu'à la mise en œuvre du problème inverse en exploitant des propriétés sur le spectre du signal mesuré. La contribution suivante ambitionne d'aller plus loin que la simple localisation d'activités par le problème inverse afin de donner accès à la dynamique de l'activité corticale. Partant des estimations de sources sur le maillage cortical, la méthode proposée utilise des méthodes d'optimisation combinatoires à base de coupes de graphes afin d'effectuer de façon robuste le suivi de l'activité au cours du temps. La dernière contribution de cette thèse porte sur l'estimation de paramètres sur des données M/EEG brutes non moyennées. Compte tenu du faible rapport signal à bruit, l'analyse de données M/EEG dites "simple essai" est un problème particulièrement difficile dont l'intérêt est fondamental afin d'aller plus loin que l'analyse de données moyennées en explorant la variabilité inter-essais. La méthode proposée utilise des outils récents à base de graphes. Elle garantit des optimisations globales et s'affranchit de problèmes classiques tels que l'initialisation des paramètres ou l'utilisation du signal moyenné dans l'estimation. L'ensemble des méthodes développées durant cette thèse ont été utilisées sur des données M/EEG réels afin de garantir leur pertinence dans le contexte expérimental parfois complexe des signaux réelles M/EEG. Les implémentations et les données nécessaires à la re-

productibilité des résultats sont disponibles. Le projet de rétinotopie par l'exploitation de données de MEG a été mené en collaboration avec l'équipe du LENA au sein de l'hôpital de La Pitié-Salpêtrière (Paris).

Mots clés:
neuroimagerie, magnétoencéphalographie (MEG), électroencéphalographie (EEG), vision, rétinotopie, méthode des éléments frontières, problème inverse, optimization convexe, regression parcimonieuse, analyse simple-essai, coupes de graphes.

ACKNOWLEDGMENTS

First I would like to express my deep gratitude to Maureen Clerc for supervising my work and sharing her expertise during these three years. I would also like to thank Olivier Faugeras for having welcomed me in his research team and giving me the opportunity to work in two prestigious and stimulating institutes: the INRIA Sophia Antipolis and the Ecole Normale Supérieure in Paris.

I would also like to thank Matti Hämäläinen and Rémi Gribonval for reviewing my thesis and taking the time to make insightful remarks and suggestions on my work. I am also extremely grateful to Sylvain Baillet, Eric Moulines and Elsa Angelini for their participation in the jury.

Many people have contributed to the work presented in this thesis and I want to thank them here: Théo Papadopoulo for helping me clarifying my thoughts on many subjects and sharing his "geeky side", Francis Bach for always being available and sharing his expertise on optimization and machine learning with so much simplicity, Sylvain Baillet for mentoring me in the MEG community, Renaud Keriven for introducing me to combinatorial optimization, Rachid Deriche for being my new INRIA team leader and helping me with my doubts and interrogations, Sylvain Arlot for sharing his knowledge on non-parametric statistics, Jean-Yves Audibert for his feed back on manifold learning methods, Bertrand Thirion for the past but mostly the future, Jean Lorenceau for assisting me in the hard moments of experimental data acquisition, Benoit Cottereau for challenging me when comes the problem of analyzing real data, Matthieu Kowalski for our fruitful collaboration, Demian Wassermann for putting up with me in his office, Nicole Voges for proofreading part of this thesis, Gabriel Peyré for always answering my questions, Stanley Durrleman for all the hours we spent together in the traffic talking about our research and Marie-Cécile Lafont for all her help.

I will not forget all the friends I have made during these three years: Sylvain Vallaghe, Florence Gombert, Julien Lefèvre, Maxime Descoteaux, Jonathan Touboul, Maria-José Escobar, Emmanuel Caruyer, Auro Ghosh, François Grimbert, Adrien Wohrer, Romain Veltz, Mathieu Galtier, Joan Fruitet, Emmanuel Olivi, Rodolphe Jenatton and the "dream team" at the ENS: Michael Péchaud, Pierre Maurel and Patrick Labatut.

I would also like to thank my parents for all the efforts they made for me since day one.

Finally, I would like to thank Claire for her love and support day-to-day during these years.

Contents

List of Tables

List of Figures

List of Abbreviations

AP	Action potential
BEM	Boundary Element Method
CNS	central nervous system
ECD	equivalent current dipole
EEG	Electroencephaphy
FDM	Finite Difference Method
FDR	False Discovery Rate
FEM	Finite Element Method
fMRI	Functional Magnetic Resonance Imaging
GCV	generalized cross-validation
M1	Primary motor cortex
MEG	Magnetoencephalography
MN	Minimum-Norm
nIRS	Near-infrared Spectroscopy
PET	Positron Emission Tomography
PLV	Phase locking value
PSP	Post-synaptic potential
ReML	Restricted Maximum Likelihood
S1	Primary somato-sensory cortex
SBL	Sparse Bayesian Learning
SNR	Signal to noise ratio
SVD	Singular Value Decomposition
SymBEM	Symmetric Boundary Element Method
TV	Total variation
V1	Primary visual cortex
WMN	Weighted Minimum-Norm

Introduction

CONTEXT

With approximately 10^{12} neurons in the central nervous system (CNS), 10^{15} synaptic connections releasing and absorbing 10^{18} neuro-transmitter and neuro-modulator molecules per second, the human brain is an object of prodigious complexity. If it were a computer, it would be capable to process 10^{12} Gigabits of information per second, all in about 1.6 Kg of weight and with a power consumption of 10-15 Watts [68]. The study of the brain activity with medical imaging methods is named *functional neuroimaging*.

In the last 30 years, neuroimaging has been a very active field of research. From 1985 to 2005, the number of related publications has increased by an order of magnitude. Functional neuroimaging however has a history that dates back far earlier. Human brain activity was first recorded by Hans Berger (1929) [98] who measured the first electroencephalogram (EEG) in humans. Later, in the 60's and 80's, several other neuroimaging techniques were introduced. The best known are magnetoencephalography (MEG), positron emission tomography (PET), functional magnetic resonance imaging (fMRI) and near-infrared spectroscopy (nIRS). fMRI is the most popular functional neuroimaging modality. One reason is that MRI scanners used for anatomical imaging can also be employed for functional imaging. EEG with its cheap instrumentation cost comes next, followed by PET, MEG, and nIRS. Even if fMRI is the most popular modality in the neuroimaging field, statistics prove that MEG and EEG research received a growing interest in the 90's, which can be explained by the improvement of acquisition devices, by the development of MRI as an anatomical basis for M/EEG studies, and also by the development of new methods adapted from other research fields such as signal and image processing, statistics, and scientific computing.

MEG and EEG (collectively M/EEG) are electromagnetic brain imaging modalities whose interest comes from the electric nature of neuronal communications. Neurons communicate with the displacements of electric charges that produce tiny currents. Neurons can be seen as tiny current generators. In order to produce electromagnetic fields detectable outside of the head, multiple neurons within a same structure need to act in concert. As opposed to fMRI that measures differences of blood oxygenation associated with the neuronal activations, M/EEG have a direct and instantaneous access to the electric phenomena. Therefore, MEG and EEG have an excellent temporal resolution.

In order to measure the electric potentials generated by neuronal activity an EEG device consists of a set of electrodes that are applied on the scalp so to establish electrical contact with the skin. Modern full head EEG caps can have nowadays more than a 200 electrodes. The counterpart of EEG is MEG that measures the magnetic fields generated by the neuronal activity. The first MEG measurements date back to the research of David Cohen in the 60's [35]; the first whole head MEG systems with hundreds of sensors capable of imaging the entire brain became available in the early 90's.

Traditionally, the EEG analysis has been based on inspection of the morphology of waveforms. As a matter of fact, most current practice of EEG in neurology is still based on these first attempts. Typically, neurological clinics perform EEG examinations for epilepsy, sleep disorder, migraine, and a few other pathological conditions for which the waveform bears diagnostic utility, as for spikes, spindles, generalized slowing, temporal theta, *etc*. Meanwhile, basic electrophysiological research has taken a different path.

The development of digital computers, together with the advances of signal processing methods contributed to transform M/EEG data analysis into a domain of research for engineers, physicists and mathematicians. Assisted by the invention of the FFT algorithm in 1965 by Cooley and Tuckey, frequency-domain analysis of EEG time series, such as power spectral density estimation or phase coherence, has been used since the 60's for cognitive and clinical studies. Time-frequency analysis of event-related synchronization and desyn-

chronization (ERS/ERD) has provided means to study brain dynamics in the scale of tens of milliseconds, preserving both spatial and spectral information. This has extended task-related brain studies beyond evoked response potentials and morphology. More recently, in the 90's, the advances of anatomical MRI data, giving access to individual brain anatomy, marked the transition into the era of functional localization of M/EEG activity.

Results of functional localization of M/EEG activity can be seen as 3D volumetric or 2D surface images of the living brain. While a standard movie is streamed with 25 images per second, accurate functional imaging with M/EEG could provide around a thousand images of the brain per second. However accurate functional localization of M/EEG activity with a temporal resolution of 1 kHz is a partially solved problem and is still a major challenge of M/EEG data analysis. To reach this goal, various computational and mathematical challenges need to be tackled, turning the study of the brain activations with M/EEG in a strongly multidisciplinary field of research at the crossroads of neurophysiology, signal processing, electromagnetism, multivariate statistics, and scientific computing.

In this thesis, various mathematical and computational aspects of M/EEG data analysis are covered, with the constant objective of being able to achieve accurate localization in space and reconstruction of the dynamics of neural activity. Our contributions start by the accurate modeling of the head as a medium that propagates the electromagnetic fields produced by the neurons. This problem, known as the forward problem, has a unique solution. The solution can be obtained analytically for spherical head models but requires numerical solvers when realistic head models are considered. Improving the speed and accuracy of such solvers, but also facilitating their usability in the M/EEG community is the topic covered by the first part of this thesis.

To estimate the current generators underlying noisy M/EEG data, one has to solve an electromagnetic inverse problem. Theoretically, a specific electromagnetic field pattern may be generated by an infinite number of current distributions. Fortunately, physiological and anatomical information can be employed to constrain the solution. The problem is said to be ill-posed. In this thesis we focus on distributed inverse solvers. The use of such solvers is motivated by their ability to provide localization results for activation patterns involving multiple generators distributed over the entire brain. In order to tackle this challenging problem, we provide in this thesis very efficient optimization methods in order to get algorithms tractable on real datasets. Our methodological contributions go beyond the inverse problem by proposing a method to robustly follow over time the activations on the cortex. The main motivation for the development of the methods detailed in this thesis, was to contribute to the study of human vision with MEG and EEG. Throughout this thesis, methods are tested with real M/EEG data in order to prove their effectiveness and relevance for clinical and cognitive M/EEG studies.

ORGANIZATION AND CONTRIBUTIONS OF THIS THESIS

Chapter 1 - Neural basis of EEG and MEG
The EEG and MEG signals are generated by the electrical activity of the neurons. At the cellular level, displacements of electric charges create tiny differences of potential. In the cortex, groups of neurons, particularly pyramidal neurons in the cortex, form structured assemblies that, when simultaneously active, produce electromagnetic fields detectable outside of the head. Human EEG recordings date back to 1929 with the German physiologist and psychiatrist Hans Berger, while the first MEG recordings were obtained in the late 60's by David Cohen. In this first chapter, we review the physiological basis of the generation of the signal measured by MEG and EEG and provide some details on the evolution of acquisition devices from their discovery to the most recent systems.

Chapter 2 - The forward problem
Understanding how a current generator located inside the head can produce a distribution of potential on the scalp or a magnetic field outside of the head is called the *forward problem*. Because of the low frequency of the signals measured with M/EEG, the time derivatives in the Maxwell's equations can be neglected. In this quasi-static approximation, the forward modeling implies that the signal measured on the sensors is the instantaneous sum of the signals produced by each current generator. However, computing this linear operator, *i.e.,* solving the forward problem with a realistic head model can be mathematically and computationally challenging. In this chapter, we review existing methods to solve the forward problem with different assumptions for the conductor geometry of the head. With realistic head models the solution is not known analytically and is obtained with numerical solvers. The first contribution of this thesis is on the efficient and precise numerical resolution of this problem using a Boundary Element Method (BEM) called the Symmetric BEM. This approach is compared to alternative open source solvers, demonstrating its excellent precision.

Chapter 3 - The inverse problem with distributed source models
While the forward problem provides the link between the measured signal and the neural current generators, the *inverse problem* aims at estimating the positions and amplitudes of these generators from a limited number of noisy measurements. Three types of approaches exist: parametric methods also referred as dipole fitting, scanning techniques and image-based methods with distributed source models. The latter approach formulates the inverse problem as a deconvolution problem where the convolution operator, or smoothing kernel, is the solution of the forward problem. Such an approach offers a rigorous formulation of the inverse problem without making strong modeling assumptions. However, the problem is strongly ill-posed. The solution of such problems classically requires to impose constraints or priors on the solution. This chapter is dedicated to the presentation of priors based on the ℓ_2 norm. Implementation details and practical information are carefully detailed. The presentation covers standard minimum-norm methods, noise normalized solutions (dSPM and sLORETA), spatio-temporal solvers, and finally Bayesian approaches where the prior is not fixed *a priori* but learned from the data.

Chapter 4 - M/EEG inverse modeling with non differentiable constraints and sparse priors
Standard ℓ_2 priors lead to very convenient linear inverse solvers but produce source estimates smeared out over the cortex. The ℓ_2 prior is said to lead to solutions with high diversity, as opposed to solutions with high sparsity where only a few sources have non-zero activations.

Such a behavior of the ℓ_2 norm can become problematic when one attempts to achieve precise localization of focal sources. In order to reduce this problem, Bayesian learning of the prior can be an alternative. In this chapter, we investigate priors where the sparsity of the reconstruction is induced by the choice of the prior. The ℓ_1 norm has this interesting property and has proved its ability to efficiently solve very challenging ill-posed problems in signal processing and machine learning. Unfortunately, such a prior leads to non differentiable optimization problems for which the solutions cannot be obtained in closed-form as in the ℓ_2 case. In this chapter, we review some algorithms that can be used to efficiently solve ill-posed problems involving the ℓ_1 norm. We promote iterative algorithms based on the use of proximity operators and show that they provide a very general approach for solving inverse problems previously introduced in the M/EEG literature. We also explain how structured sparsity with mixed norms can be used to provide an efficient spatiotemporal solver and develop a new framework to compute source estimates for multiple experimental conditions simultaneously using an inter-condition prior.

Chapter 5 - Fast retinotopic mapping with MEG
This chapter presents a direct application of the previous chapters to a real case study. The objective of this study was to achieve retinotopic mapping with MEG. The motivation for this work was twofold. First, we wanted to demonstrate that MEG could reproduce the retinotopic maps obtained by standard protocols in fMRI. Second, thanks to the excellent temporal resolution of MEG, we gain access to brain dynamics during visual processing. In this chapter, we present the anatomical basis of the human visual system, detail the experimental protocol we contributed to design, and also the methodological tools we implemented in order to provide retinotopic maps with MEG. The protocol is based on steady-state visual evoked potentials. We discuss the algorithmic details of the signal extraction procedure and our method for non-parametric statistical tests. We present results obtained with linear inverse solvers and illustrate their limitations. To address these limitations, we propose to include all the experimental conditions simultaneously in the analysis and to use an inter-condition sparse prior based a mixed norm described in the previous chapter. Finally, we give some insight on how timings and delays of propagation could be extracted from the phase of the Fourier spectrum of the source activation time series.

Chapter 6 - Tracking cortical activations with graph cuts
The work presented in this chapter attempts to go one step further from source localization in order to provide a clear representation of the cortical dynamics during neural processing. The linear ℓ_2 inverse solvers are convenient to use but produce huge amounts of data out of which the relevant information needs to be extracted. The purpose of our contribution presented in this chapter, is to extract from the mass of data provided by distributed inverse solvers the spatio-temporally consistent activations. The algorithm provides a robust and principled way to track the "hot-spots", *i.e.,* active regions, over the triangulated cortical mesh. A variational formulation of the problem is derived and a very efficient optimization method based on graph-cuts is detailed in order to find globally optimal solutions.

Chapter 7 - Graph-based estimation of 1-D variability in event related neural responses
The last contribution of this thesis addresses a particularly challenging problem in M/EEG data processing: parameter estimation from single-trial data. In classical M/EEG data processing pipelines, the signal-to-noise ratio of the measured data is improved by averaging multiple recordings obtained under the same experimental conditions. By doing so, one assumes that the signal of interest is the same in each repetition, also called a trial. This is

unfortunately not true, as the neural response of the subject can vary, typically because of habituation effects, anticipation strategies, or fatigue. This is particularly the case for brain responses occurring late after the stimulation. Such late activations can correspond to higher cognitive levels of processing and are therefore of major interest to better understand how our brain performs complex cognitive tasks. The method uses advanced graph-based methods and has numerous advantages over alternative strategies: trial averaging is not used in the estimation, they provide solutions with global optimality, thus avoiding initialisation problems, finally thanks to the efficiency of the method, parameters can be rapidly estimated by cross-validation and grid search.

Appendices

Appendix A - Kronecker products
This appendix is a brief introduction to the manipulation of Kronecker products. The Kronecker product is a valuable tool to manipulate spatiotemporal regularizations as illustrated in chapter 3.

Appendix B - Introduction to graph cuts
In this appendix, we present the basic concepts on graph cuts in order to facilitate the understanding of the optimization methods used in chapter 6 and chapter 7.

Appendix C - Time frequency analysis with Gabor filters
This appendix contains a description of the Gabor filters used to compute the time-frequency analysis results presented in chapter 5.

Appendix D - Publications of the author
In this appendix, we list the submitted and the already published material from the author.

Software contributions

Finally, we would like to point out that all the algorithms presented in this thesis are available on the INRIA Forge.

The forward solver OpenMEEG detailed in chapter 2 is available at:

`https://gforge.inria.fr/projects/openmeeg/`

The Matlab interface we developed was integrated into the current release of Fieldtrip and is available for download from the Fieldtrip home page:

`http://fieldtrip.fcdonders.nl/`

All the implementations of the inverse solvers presented in chapters 3 and 4, with also the code to perform the tracking detailed in chapter 6, are available in a MATLAB Toolbox called EMBAL (Electro-Magnetic Brain Activity Localization):

`https://gforge.inria.fr/projects/embal`

Most of the figures presented in this thesis are done with the functions implemented in EMBAL.

Finally the EEGLAB plugin to perform parameter estimation on single-trial M/EEG data as described in chapter 7 is available here:

`https://gforge.inria.fr/projects/eeglab-plugins/`

CHAPTER 1

NEURAL BASIS OF EEG AND MEG

MEG and EEG measure the electromagnetic signal produced by the activity of our brain. To provide more insight into the physiological phenomena behind M/EEG measurements, this first chapter discusses the biological aspects of the functioning of the human brain.

Contents

1.1 ANATOMY AND ELECTROPHYSIOLOGY OF THE HUMAN BRAIN

1.1.1 General brain structures: From macro to nano

Together with the spinal cord, the brain forms the central nervous system (CNS). It is the largest part of the nervous system and is itself composed of a lower part, the brainstem, and an upper part, the prosencephalon, a.k.a., the forebrain. In figure 1.1 the brainstem includes the mesencephalon, the medulla, and the pons. It connects the two remaining structures that form the prosencephalon, i.e, the telencephalon and the diencephalon, to the spinal cord. The medulla, or lower part of the brain stem, controls unconscious activity of muscles and glands involved in breathing, heart contraction, salivation, *etc*. Just above the medulla, the pons connects the two hemispheres of the cerebellum which is located in the inferior posterior portion of the head (directly dorsal to the pons). The diencephalon is located in the midline of the brain and contains the thalamus and the hypothalamus. The most superior structure, the telencephalon, or cerebrum, includes the lateral ventricles, the basal ganglia and the cerebral cortex.

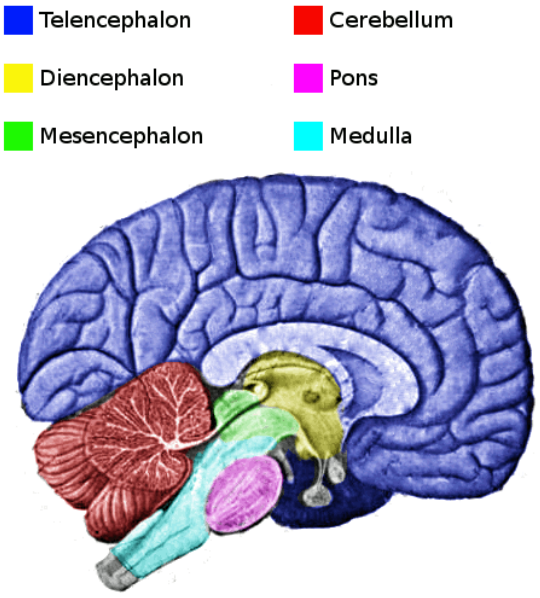


Figure 1.1: Main anatomical structures of the vertebrate brain (Source wikipedia.org).

An axial slice (see figure 1.3 for naming conventions) of the cerebrum presented in figure 1.2 exhibits two main structures: the white matter and the gray matter. The gray matter of the cerebrum forms the cerebral cortex, a.k.a., the neocortex. The neocortex forms the majority of the cerebrum and corresponds to its most exterior part. It has a left and a right hemisphere (see figure 1.4). It is assumed that the neocortex is a recently evolved structure, and is associated with "higher" information processing by more fully evolved animals (such as humans, primates, dolphins, *etc*.).

Each hemisphere of the neocortex is generally divided into 4 lobes as represented in figure 1.5.

The following functions can be roughly related to each lobe:

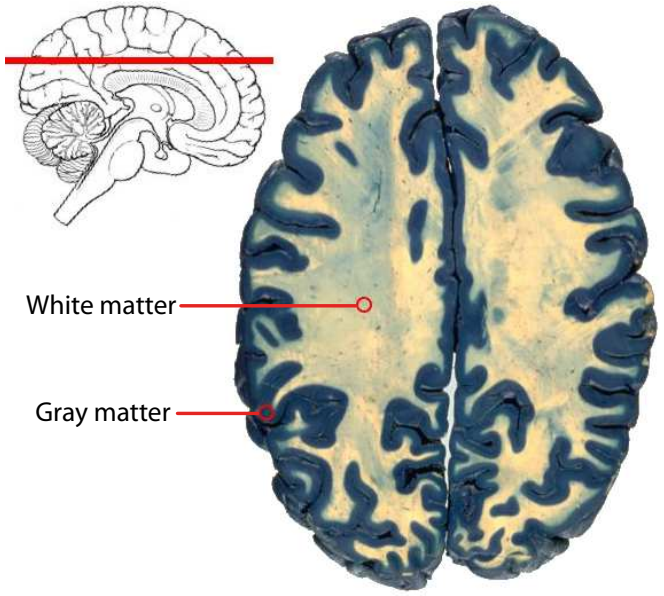


Figure 1.2: Axial slide of the brain (Adapted from: dartmouth.edu).

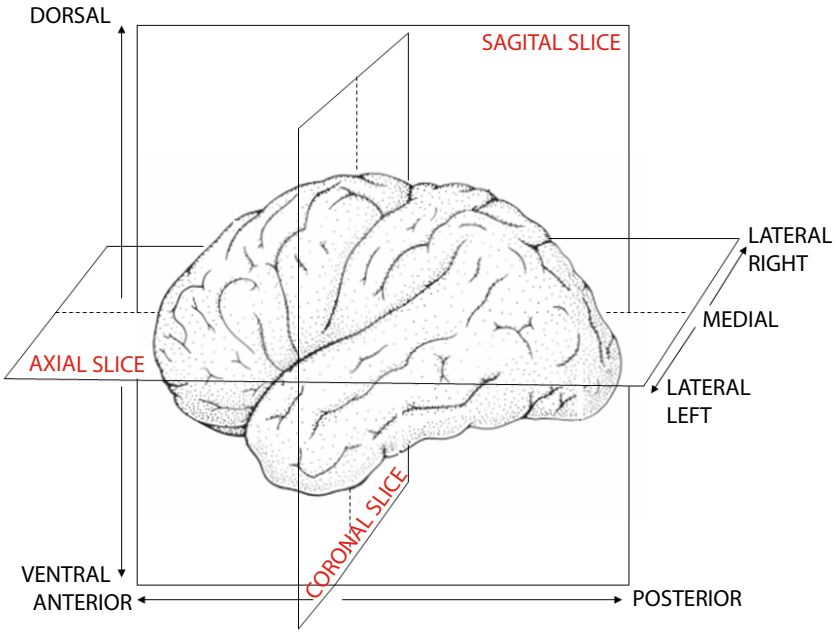


Figure 1.3: Standard naming conventions for planar slices through the brain.

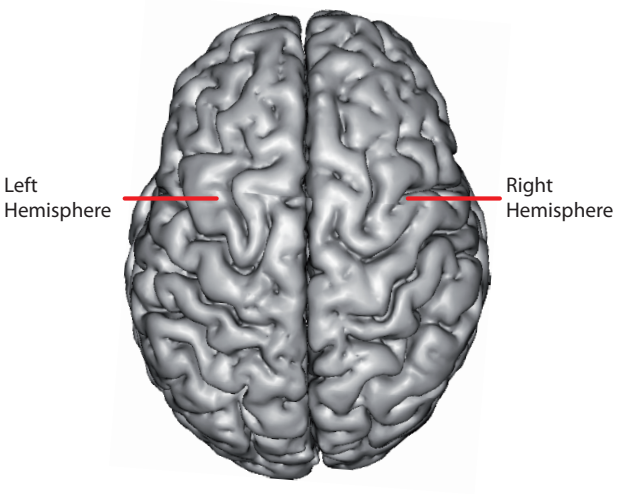


Figure 1.4: Brain hemispheres. At first glance the two hemispheres are very similar but their detailed structure is clearly different.

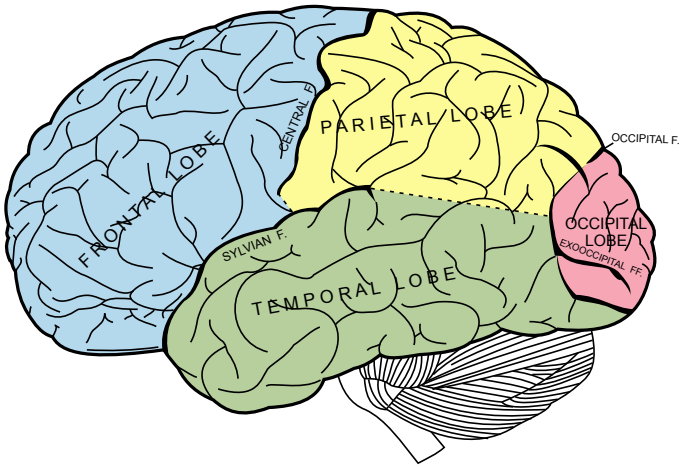


Figure 1.5: The different lobes of the cerebral cortex: the occipital, the parietal, the temporal, and the frontal lobes (From 20th U.S. edition of Gray's Anatomy of the Human Body, 1918 (public domain)).

- Frontal Lobe: associated with reasoning, planning, parts of speech, movement, emotions, and problem solving.
- Parietal Lobe: associated with movement, orientation, recognition, perception of stimuli, and speech.
- Occipital Lobe: associated with visual processing.
- Temporal Lobe: associated with perception and recognition of auditory stimuli, memory, and speech.

Lobes are separated by major fissures that are present in all individuals. This makes the identification of the different lobes on a particular subject possible by simple visual inspection. For example, the parietal and the frontal lobes are separated by the central fissure, a.k.a. the central sulcus, and the temporal lobe is separated from the parietal lobe and the frontal by the Sylvian fissure (cf. figure 1.5). Fissures are also commonly called sulci.

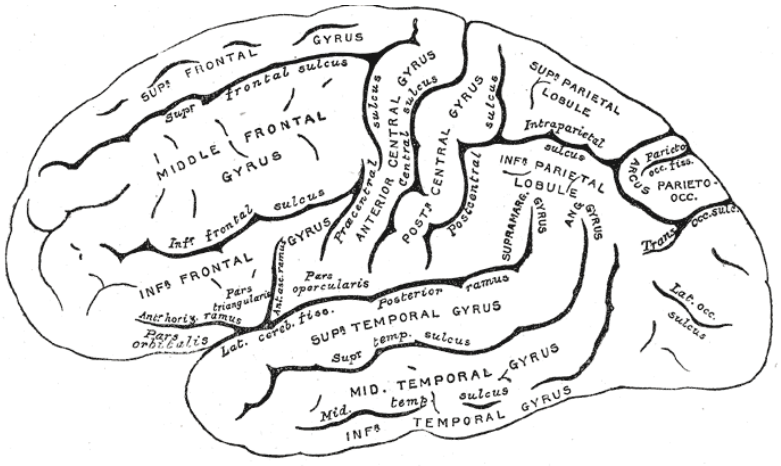
The counterpart of the cortical fissures are the gyri. Gyri are the structures between the fissures. The main gyri are presented in lateral and medial views in figure 1.6. Some of the gyri contain brain regions with known cognitive functions like the post-central gyrus that includes the primary somatosensory cortex (S1), *cf.,* figure 1.7.

Such a knowledge on the localization of some brain functions is particularly interesting from a methodological point of view as it provides a way to achieve validation. Many M/EEG methodological tools are tested on datasets involving somesthetic stimulation. This is the case, for example, in chapter 4 and chapter 6.

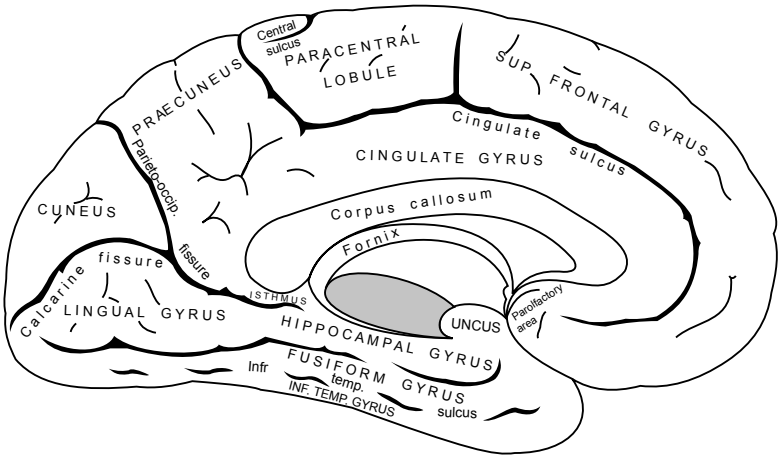
A closer look at the gray matter shows that its structure varies across the different regions. The structural properties of the gray matter include the number of layers (see figure 1.8), the cell composition, the thickness and organization. These properties, called by neuroanatomists cytoarchitectonic properties, are not the same over the whole surface of the cortex. Their differences led, in 1909, the neuroanatomist Korbinian Brodmann to divide the cortex into regions called Brodmann areas (see figure 1.9) whose historical characteristics were homogeneous [25]. Some functions were then assigned to some of these areas. For example the visual cortex, which is the object of an MEG study in chapter 5, corresponds to areas 17 and 18. Although, even if this subdivision of the cortex in Brodmann areas seems very convenient, its utility in brain functional imaging studies are usually limited labelling a particular brain region: it is more convenient to write Brodmann area 5 (BA5) than the "posterior part of the post-central gyrus".

Generally speaking, most of the cortex is made up of six layers of neurons, from layer I at the surface of the cortex to layer VI, close to the white matter. For humans, the cortical thickness varies from 3 to 6 mm. The organization of the cortex is not only laminar. It has been observed that neurons one meets when moving perpendicular to the cortex tend to be connected to each other and to respond to precise stimulations with similar activities throughout the layers. They form a *cortical column*. This columnar organization of the cortex was discovered by Mountcastle with a pioneering experiment in 1957 [157]. With electrode recordings, he showed that neurons inside columns of 300 to 500 μm of diameter displayed similar activities. This is illustrated in figure 1.10. More detailed information about cortical structure and function can be found in [121, 124, 175].

The gray matter is composed of neurons and glial cells. The human brain contains around 10^12 neurons. The neurons are linked together and each neuron has up to 10000 connections. The neuron is a cell with a special shape: it is composed of a soma or cell body, containing the nucleus, a dendritic tree and an axon, as shown in figure 1.11. The white matter is formed predominantly by myelinated axons interconnecting different regions of the central nervous system.



(a) Gyri lateral view



(b) Gyri medial view

Figure 1.6: Main gyri presented in lateral (a) and medial (b) views (From 20th U.S. edition of Gray's Anatomy of the Human Body, 1918 (public domain)).

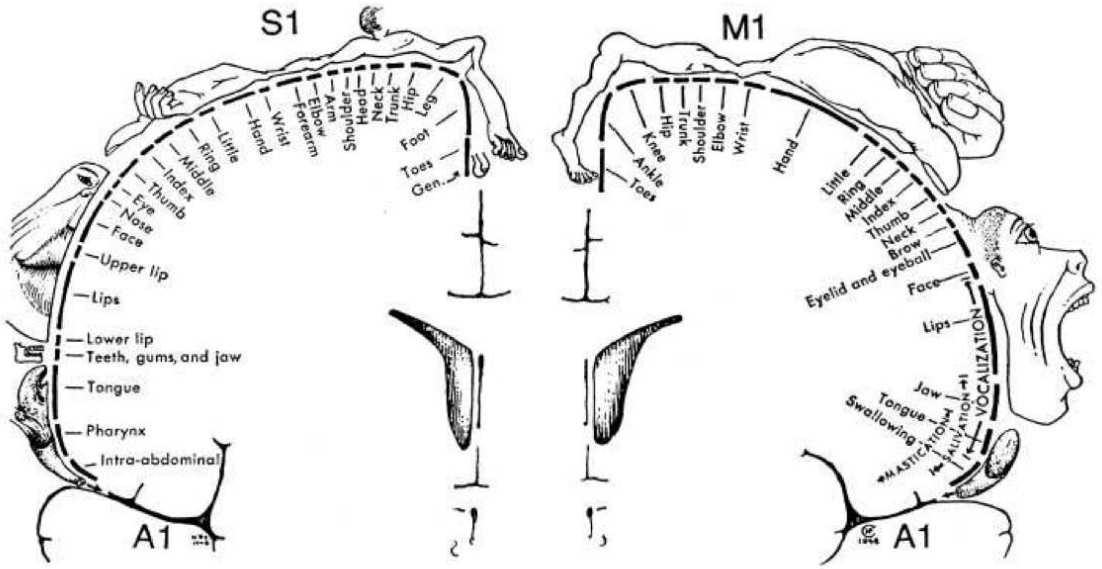


Figure 1.7: Cortical homunculus by Wilder Graves Penfield [174]. It represents the mapping the primary sensory (S1) and primary motor (M1) cortex. S1 lies on the posterior wall of the central sulcus (cf. post central gyrus in figure 1.6(a)) and M1 on the anterior part. These maps were established by direct electrical stimulation on patients during surgery. Primary auditory cortices (A1), left and right, are represented in the temporal lobes.

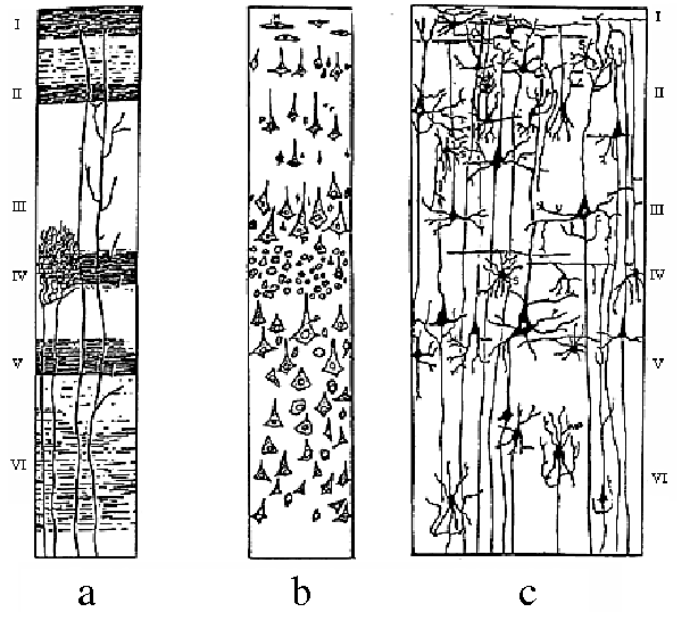
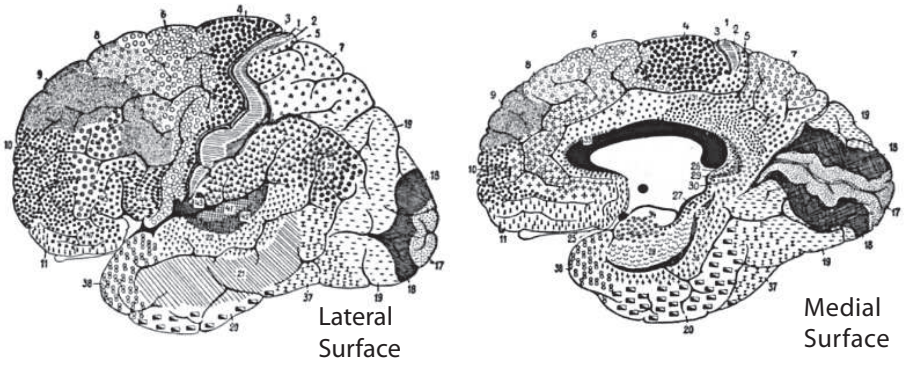
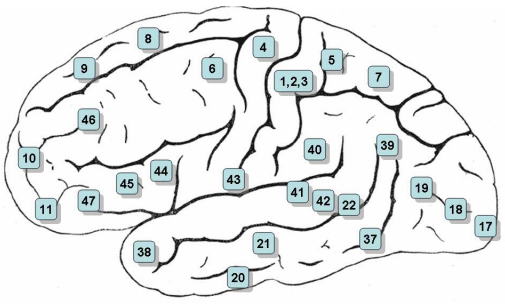


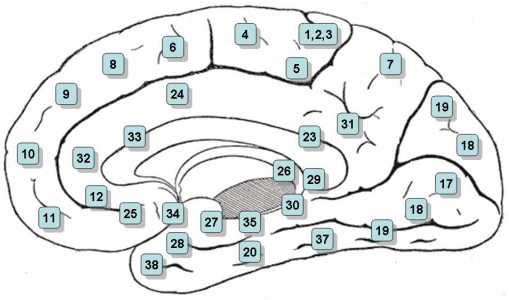
Figure 1.8: Cortical layers. Layer organization of the cortex (a) Weigert's coloration shows myelinated fibers (axons) and so the connections inside and between layers, (b) Nissl's coloration only reveals cell bodies (c) Golgi's coloration shows the whole cells (From [163]).



(a) Reproduction from original



(b) Lateral schematic view (From 20th U.S. edition of Gray's Anatomy of the Human Body, 1918 (public domain))



(c) Medial schematic view (From 20th U.S. edition of Gray's Anatomy of the Human Body, 1918 (public domain))

Figure 1.9: Brodmann areas. In 1909, Brodmann [25] divided the cortex into 52 cytoarchitectonic areas according to the thickness of the cortical layers. For example, layer IV is very thin in the primary motor cortex (area 4) while it is very thick in the primary visual cortex (area 17).

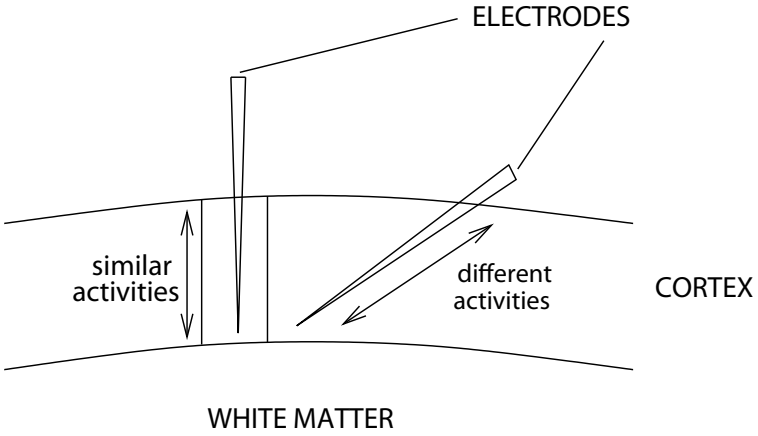


Figure 1.10: Mouncastle's experiment and the discovery of the columnar organization of the cortex. When he moved an electrode perpendicular to the cortex surface, he encountered neurons with similar electrical activities while moving the electrode obliquely gave him different types of recordings. So he showed the existence of 300-500 μm wide columns in the cortex.

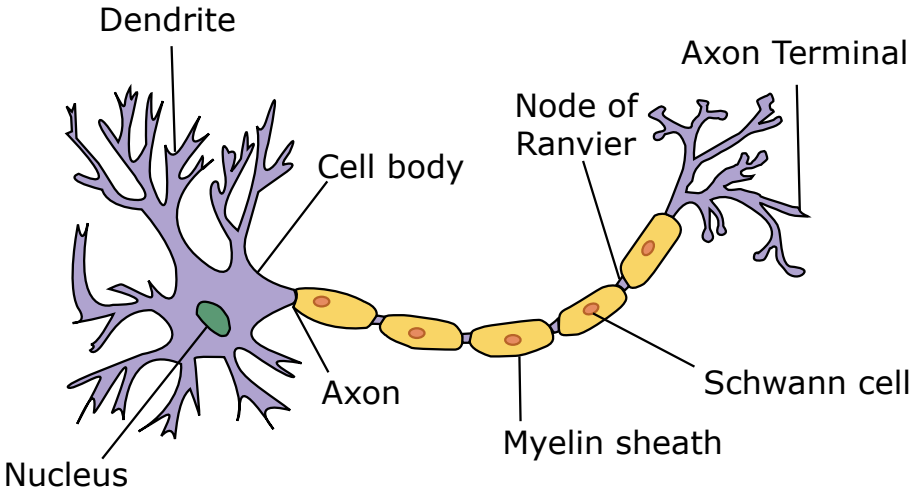


Figure 1.11: Diagram of a neuron (Source wikipedia.org)..

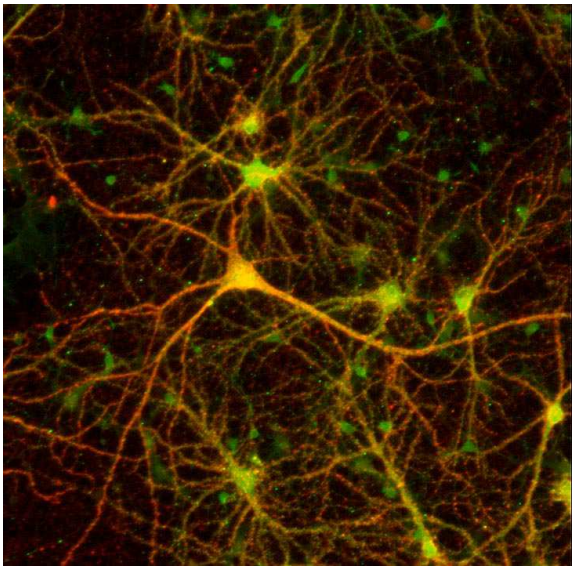


Figure 1.12: Neurons observed with an electron microscope.

1.1.2 How neurons produce electromagnetic fields

A neuron can be viewed as a signal receiver, processor and transmitter: the signal coming from the dendrites is processed at the soma and generates (or not) an action potential which is carried along the axon towards other neurons. During this process neurons produce electromagnetic fields at the basis of the M/EEG measurements.

The signals in the dendrites are called post-synaptic potentials (PSPs). The signal emitted, moving along the axon, is called the action potential (AP).

Post-synaptic potential (PSP)
The junction between the axon terminal of a neuron and a dendrite or the soma of another neuron is called a *synapse*. It can be a direct electrical junction, but synapses are mostly chemical: when an action potential reaches the end of an axon terminal, it leads to the release of neuro-transmitters. Neuro-transmitter molecules that reach an other neuron affect the membrane permeability so that specific ions (Na+ and K+) penetrate inside, increasing the resting state potential of about 10 mV with a duration of 10 ms. This is called a post-synaptic potential, shown in figure 1.13.

Action potential
If many post-synaptic potentials sum up, the membrane potential of the soma can locally reach a certain threshold which causes the neuron to "spike": some voltage-sensitive channels open, allowing positive ions to flow inside the cell, and the potential inside the neuron increases suddenly. The potential comes back rapidly to its resting state (in 1 ms), with the help of other voltage-sensitive channels that allow a compensating outward current. Because of this peak of potential, the nearby regions also reach the threshold: the action potential thus propagates along the axon, as illustrated in figure 1.14. See [125] for more details on the the ion mechanisms responsible for these two types of potentials.

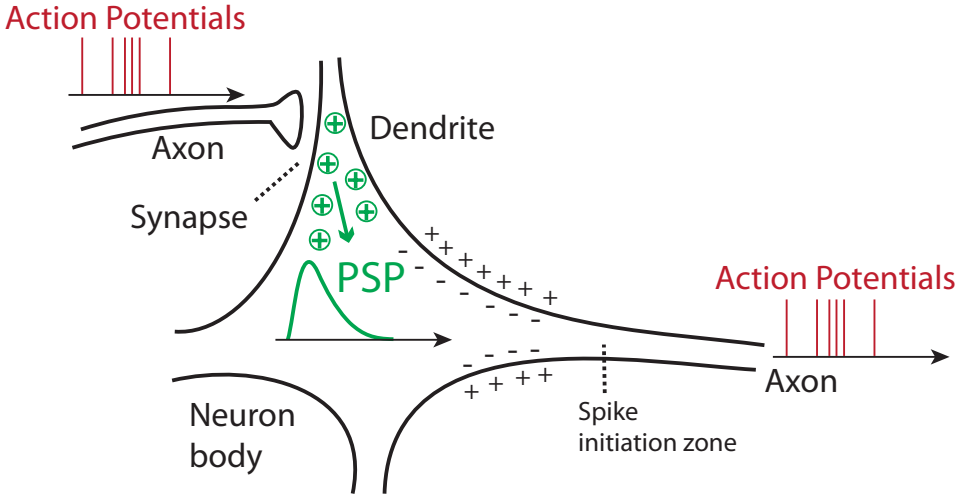


Figure 1.13: From action potentials to post-synaptic potentials (PSP). Illustration with a chemical synapse. The action potentials reach the neuron on its dendrites via chemical synapses. It creates post-synaptic potentials that by summation generate other action potentials that can propagate along the axon of the neuron.

These two types of potentials create some displacements of charges and therefore some very small currents within the neuron: the intracellular or primary currents. These currents,

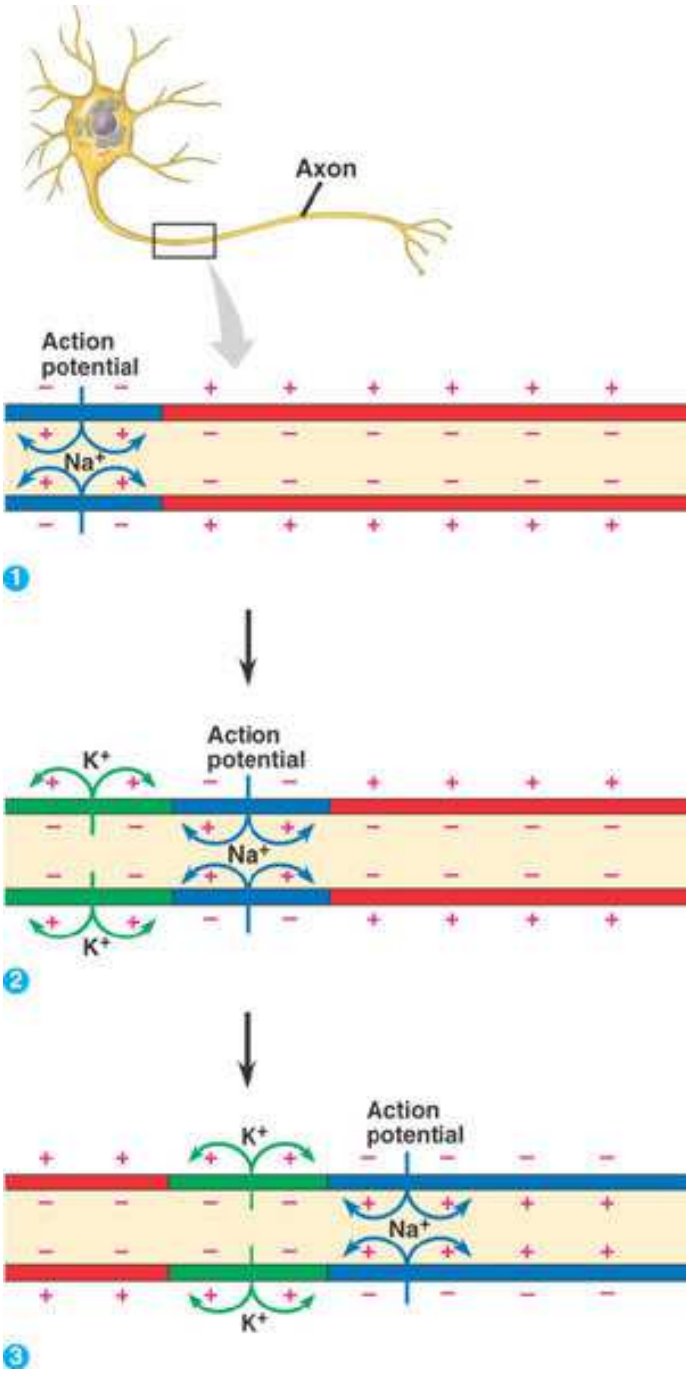
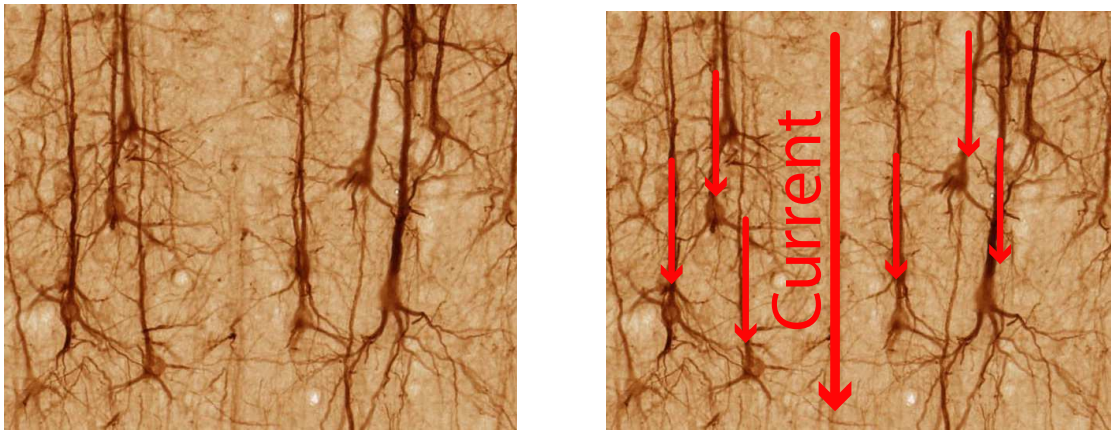


Figure 1.14: Action potential propagation (Source kvhs.nbed.nb.ca).

however, create very tiny electromagnetic fields that cannot be directly measured outside of the head with M/EEG. In order to have measurable signals, these tiny fields need to sum up. Action potentials have a temporal duration close to the millisecond making them hard to synchronize in order to sum up. On the contrary, PSPs have a temporal duration around 10 ms. This makes PSPs much better candidates to produce measurable electromagnetic fields outside the head. The temporal resolution of the phenomena points out a necessary but not sufficient condition to get good M/EEG signals. Electrical currents are vectorial quantifies. They have both an amplitude and a direction. In order to actually sum up, the currents produced by the neurons need to have a common direction. Following the conclusions of [159], it is necessary to add the field amplitudes of about 10^4 neurons with dendrites having a common direction to produce a field amplitude that is detectable from outside the head. For instance, stellate cells which have dendrites in all directions can not produce a measurable field. Only neurons called pyramidal cells have the regular geometric structure organization that is required to sum up the fields generated by their post synaptic potentials.

Pyramidal neurons

The bodies and dendrites of pyramidal neurons are located mostly in the gray matter of the cortex, and they all have a thick dendrite (called apical dendrite) extending towards the exterior of the cortex, perpendicularly to its surface, as shown in figure 1.15. These neurons constitute about 70%-80% of the neocortex, and their density is such that theoretically the simultaneous activation of an area of $1\ \mathrm{mm}^2$ of the cortex would be detectable. However, an experimental study showed that the minimal detectable activity spreads over an area of about $100\ \mathrm{mm}^2$ [100].

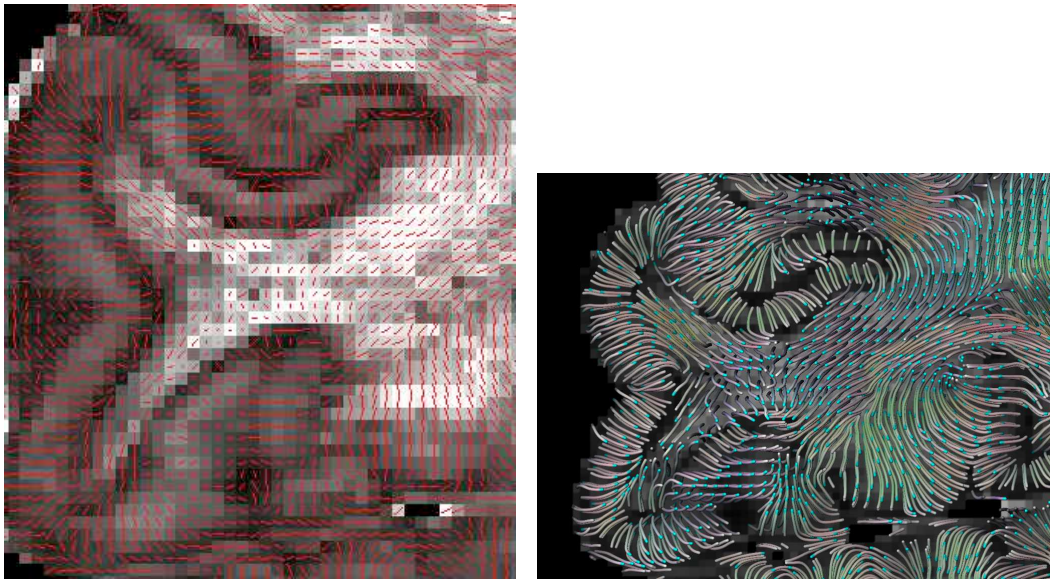


(a) Pyramidal neurons

(b) Pyramidal neurons and the produced intracellular currents.

Figure 1.15: Pyramidal neurons in medial prefrontal cortex of macaque (Source brainmaps.org).

This structured organization of pyramidal cells has been discovered by invasive studies that provided experimental results like the one presented in figure 1.15(a). Nowadays, due to the progress of brain imaging devices like MRI, more precisely diffusion MRI, this organization can be observed non invasively. Diffusion MRI offers the possibility to measure the anisotropy of the diffusion of water molecules in living tissues. This is presented in figure 1.16. In order to obtain images with such a good signal-to-noise ratio, the acquisition was performed ex-vivo with a very long period of scanning. Note that the principal directions of diffusion in the gray matter follow the organization of the cortical layers and the general structure of pyramidal neurons assemblies.



(a) Principal eigenvector directions from tensor fitting superimposed on T1 MRI

(b) Tractlets rendered in 3D (courtesy of Gordon Kindlmann)

Figure 1.16: Principal directions of water molecules diffusion estimated with tensor fitting on Diffusion MRI. Orientations appear to be very well organized with directions given by the normals to the cortical mantle. (Data: Dr J McNab & Dr K Miller, FMRIB, Oxford 3T Siemens ex-vivo whole-head diffusion imaging, .7x.7x.7mm).

Models of brain electric activity for EEG and MEG

The consequence of the latter observations for EEG and MEG is that the brain activity is observed at a macroscopic scale with respect to the size of a neuron. They capture the electrical activity of structured assemblies of neurons. The typical size of the neuron assemblies observable with EEG or MEG is larger than the size of cortical columns but smaller than the size of a cortical area. For the last three decades, neuroscientists have built models of neuron assemblies [48, 76, 113, 199, 226, 237], based on the knowledge of neuronal dynamics, but these dynamics are far from being fully understood.

These observations lead us to the problem of modeling brain electric activity for EEG and MEG. The main assumption is that the measurements corresponds to the activity of one or several assemblies of neurons. For one assembly, the EEG or MEG measurements only reflect its average activity, but usually the intrinsic dynamics of the group of neurons is unknown. As a consequence, for EEG and MEG, the most common model of the brain activity assumes that each source reflects the average activity within an assembly of neurons. The intrinsic dynamics of an assembly of neurons is hidden due to this averaging. Note that such a model agrees with the columnar organization of the cortex mentioned above. As explained in section 2.2.1, the area of a neuron assembly is small compared to the distance to the observation point (the M/EEG sensors). Therefore, the electromagnetic fields produced by an active neuron assembly at the sensor level is very similar to the fields produced by a current dipole. As a first approximation, this makes current dipoles relatively good models for active brain regions (cf. figure 1.17).

Assuming the simple dipolar model for current generators whose activity is measured by M/EEG, the electric and magnetic fields produced by an active brain region can be schematically represented like in figure 1.18 and figure 1.19. The summation of the neural currents produced by elementary generators can be approximated by an equivalent current dipole

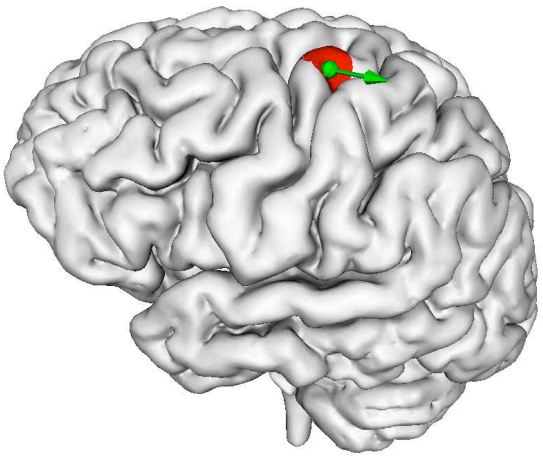


Figure 1.17: The activity of a small region of the brain can be approximated by a current dipole. The position of the dipole (the dot) is at the center of the activated cortex area (in red) and the moment of the dipole (the green arrow) corresponds to the average orientation of the pyramidal neurons in this region (perpendicular to the cortical surface).

(ECD). The electromagnetic fields produced by this ECD are strong enough to be measured outside the head. This raises the question of how to measure these fields.

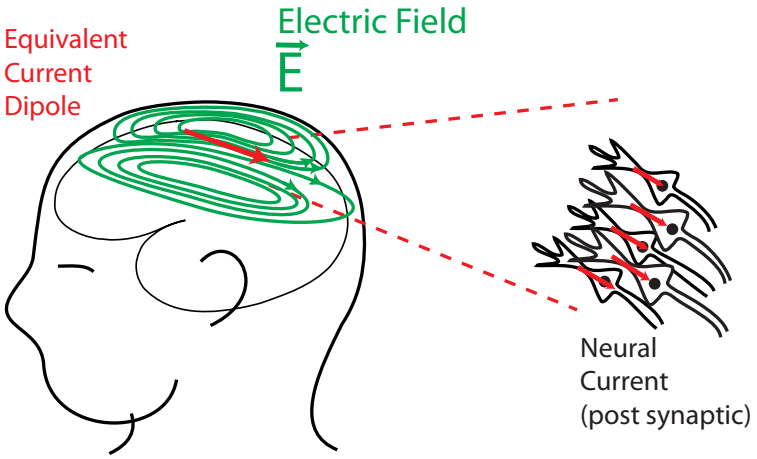


Figure 1.18: Electric field produced by neural currents modeled by an equivalent current dipole (ECD)

1.2 Instrumentation for MEG and EEG

1.2.1 Electroencephalography (EEG)

The first human EEG recordings date back to the first measurements by the German physiologist and psychiatrist Hans Berger in 1929. The recording is obtained by placing electrodes which measure the electric potential on the scalp of the subject (cf. figure 1.20).

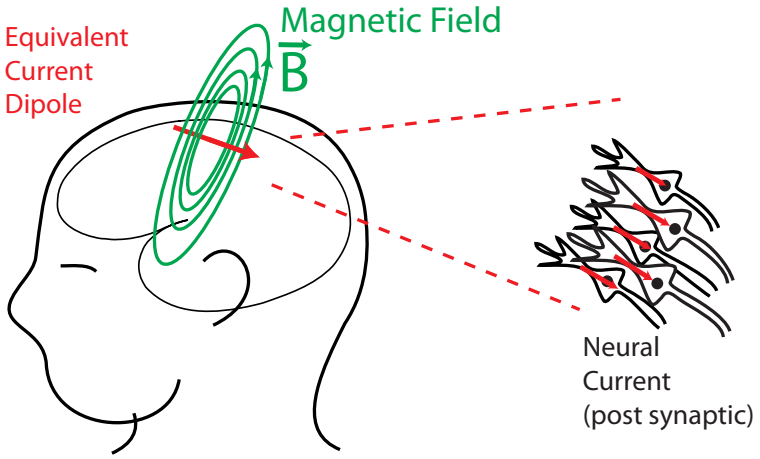


Figure 1.19: Magnetic field produced by neural currents modeled by an equivalent current dipole (ECD)

(a) EEG recordings in 1949

(b) Modern EEG recordings (Odyssée project team, INRIA Sophia Antipolis)

Figure 1.20: EEG equipment: the electrode helmet is placed on the head of the subject, then the signal is processed through an amplifier.

To obtain congruence among different laboratories, a standard electrode placement scheme was proposed by Jasper in 1958 [115], basing the positioning on head anatomical landmarks (see figure 1.22). This standardization marked the beginning of modern electroencephalography. The number of electrodes used in research has increased over the years from around 19 of Jasper's time to as many as 512 today, however the 10-20 system with 19 electrodes is still the dominant standard in clinical settings and most research is carried out with 19 to 64 electrodes.

In a modern EEG system, the electrodes are connected to an amplifier and the signals are then digitized and stored on a computer. Signals measured by EEG sensors have an order of magnitude in the range of a few μV. An example of EEG recordings is presented in figure 1.21.

The advantage of this device is its simplicity and cheap cost. Unfortunately, the low conductivity of the skull tends to diffuse the electric potential. As illustrated in figure 1.23, at the surface of the scalp, the potential only reflects roughly the underlying brain activity.

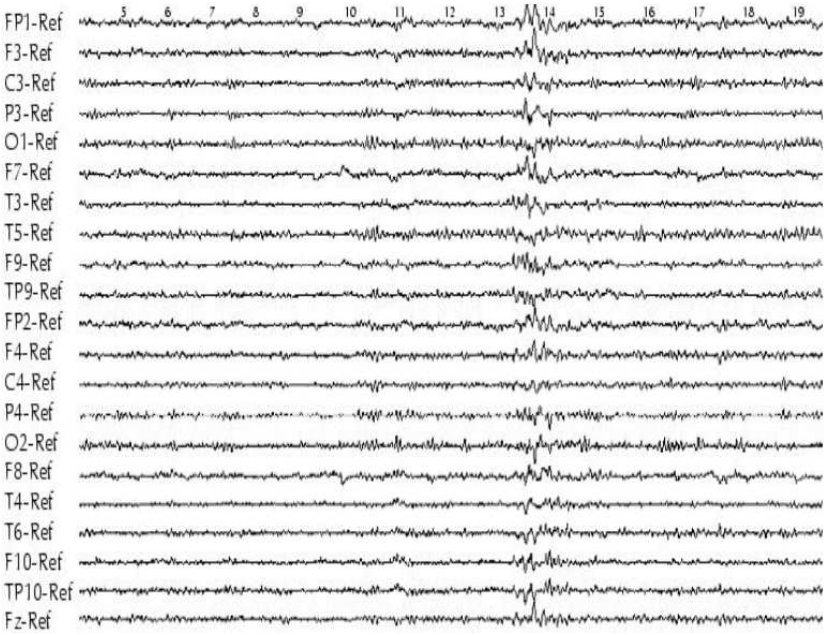


Figure 1.21: Sample EEG recordings. Each time series is the signal measured by one electrode. Electrodes have names (*e.g.,* FP1, F3, C3 *etc*.) function of their position of the scalp (cf. figure 1.22).

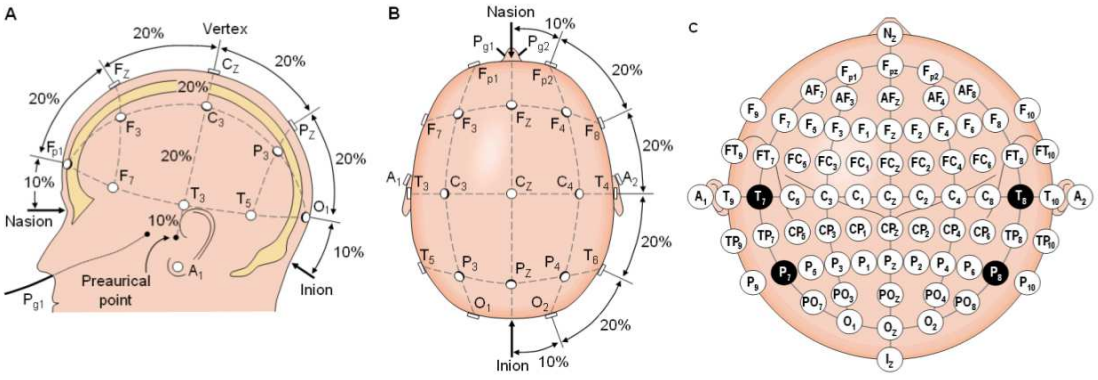
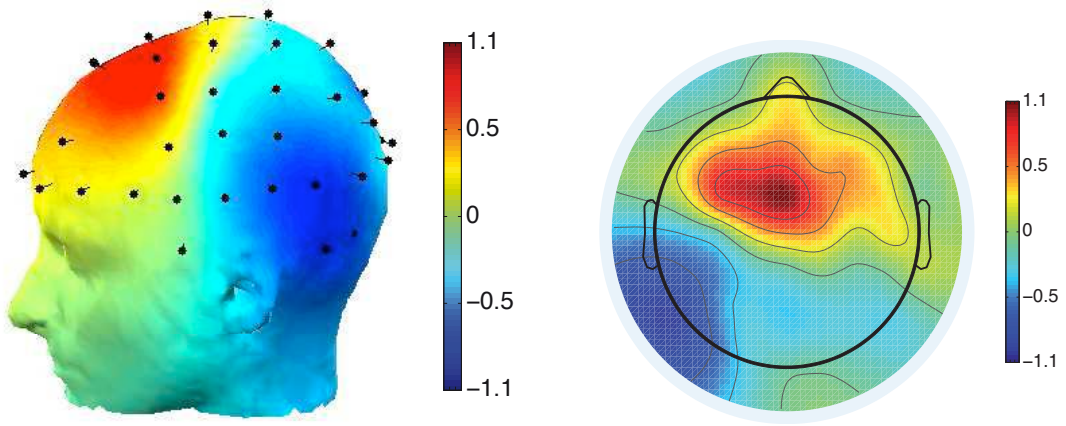


Figure 1.22: The international 10-20 system seen from (A) left and (B) above the head. A = Ear lobe, C = central, Pg = nasopharyngeal, P = parietal, F = frontal, Fp = frontal polar, O = occipital. (C) Location and nomenclature of the intermediate 10% electrodes, as standardized by the American Electroencephalographic Society. (Adapted from [67]).



(a) 3D topography (b) 2D topography

Figure 1.23: The electric potential distribution measured with EEG on a somato-sensory experiment 20 ms after stimulation (Adapted from [211]).

1.2.2 Magnetoencephalography (MEG)

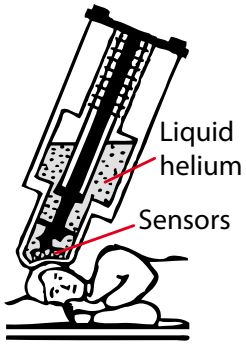
The magnetic counterpart of EEG, the magnetoencephalogram, was recorded 40 years later (1968), using room temperature coils and signal averaging on the basis of EEG [35]. Further progress in MEG required highly sensitive magnetic detectors based on superconducting and quantum phenomena and are called SQUIDs (superconducting quantum interference device). In 1969, Zimmerman and colleagues developed the first SQUIDs. They were first used for MEG in 1972 by David Cohen [36]. After this pioneering work, the field of MEG developed first by using single-channel devices, followed by somewhat larger systems with 5 to 7 channels in the mid 1980s, then systems with 20 to 40 sensor arrays in the late 1980s and early 1990s. The first MEG systems with a helmet covering the entire cortex were introduced in 1992. Today MEG systems have several hundreds channels in a helmet arrangement (see figure 1.24) allowing to capture the signal originating from the whole brain simultaneously. More details can be found in [100, 217].

MEG measurements span a frequency range from about 10 mHz to 1 kHz and field magnitudes from about 10 fT for spinal cord signals to about several pT for brain rhythms. To realize how small the MEG signals are, it should be recalled that the Earth's field magnitude is about 0.5 mT and the urban magnetic noise about 1 nT to 1 μT, which corresponds to a factor of 1 million to 1 billion larger than the MEG signals. Such large differences between signal and noise demand noise cancellation with extraordinary accuracy.

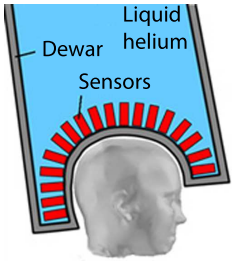
A MEG system is very expensive compared to EEG, because the SQUID sensors need to operate at very low temperature, and for this reason are immersed in liquid helium. Moreover, most often a magnetic shielded room is necessary to use the system. The main advantage of the magnetic field measurements is that it is much less sensitive to the detailed conductivity geometry of the head than the electric potential. The magnetic field observed outside the head offers a more precise representation of the underlying brain activity, see figure 1.25 in comparison to figure 1.23(b). That is why, in spite of their high cost, MEG systems are very attractive for the exploration of the human brain.

1.2.3 Other modalities for brain functional imaging

Brain functional imaging modalities can be classified in two categories: direct and indirect measures of the neuronal activity. The direct measures, like M/EEG, provide access to the



(a) Schematic representation of historical MEG device with a small number of sensors



(b) Schematic representation of full head MEG device

(c) Recent MEG device (Magnetoencephalography center, La Timone, Marseille)

Figure 1.24: MEG devices. SQUID sensors are immersed in liquid helium.

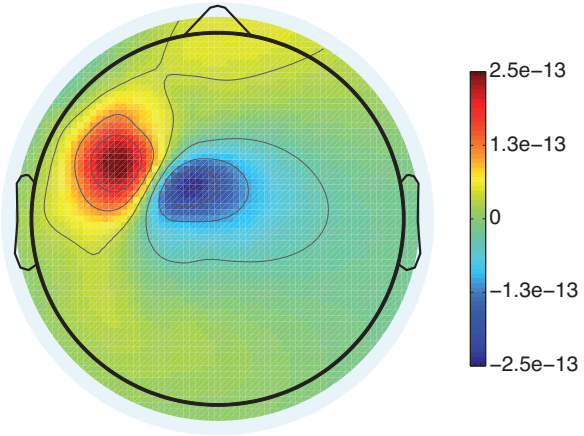


Figure 1.25: Magnetic field measured with MEG on a somato-sensory experiment. It is a 2D topography 20 ms after stimulation. Image obtained from the data used in [150].

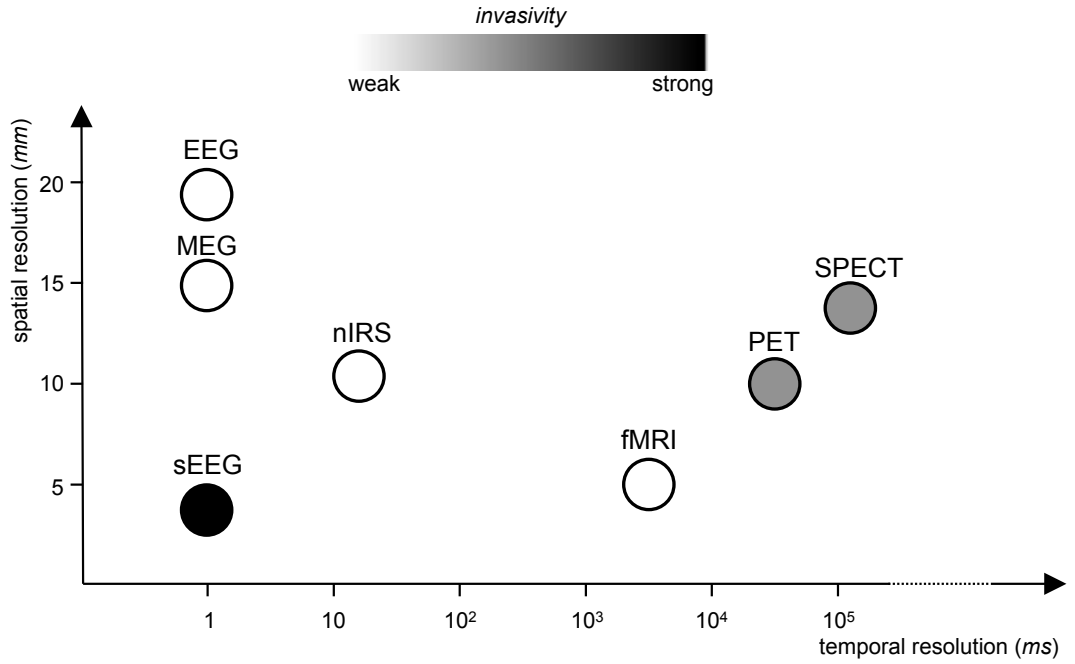


Figure 1.26: Spatiotemporal resolution and invasivity of brain functional imaging modalities.

electrical activity. Indirect measures estimate the brain activations only via the metabolic and hemodynamic processes caused by the actual neuronal activations.

Neuroimaging modalities have each some characteristic features. They can be classified in term of spatial resolution, temporal resolution and invasivity. This is summarized in figure 1.26.

Stereo-electroencephalography (sEEG)

Like M/EEG, stereoelectroencephalography (sEEG) provides access to the currents produced by the neuronal activity. By implanting depth electrodes surgically into the brain tissues, sEEG records the electrical potentials directly within the cortical layers. Electrodes are a few centimeters long and contain multiple contacts. Each contact record the local electric potential. Around the location of the activation are observed large deflections in the signal waveforms typical to sEEG recordings (cf. figure 1.27).

In the treatment of epilepsy, this ability to precisely locate the origin of a neuronal activation contributes to define the boundaries of the "epileptogenic zone", *i.e.,* the area of brain generating the epileptic seizures. It can be necessary to surgically resect this area to get rid of the epileptic seizures. This technique was introduced by the group of the Ste Anne Hospital, Paris, France, in the second half of the 20th century [111, 200].

However, this technique although has some drawbacks. The access to neuronal currents is invasive and the number of electrodes limits the recordings to very specific brain regions. In comparison, the spatial resolution of M/EEG is more limited but it records a very distributed cortical activation and is therefore not restricted to predefined brain regions.

Functional magnetic resonance imaging (fMRI)

Functional magnetic resonance imaging, or fMRI, works by detecting the changes in blood oxygenation and flow that occur in response to neural activity. An active brain area consumes more oxygen. To meet this increased demand, blood flow in the active area increases.

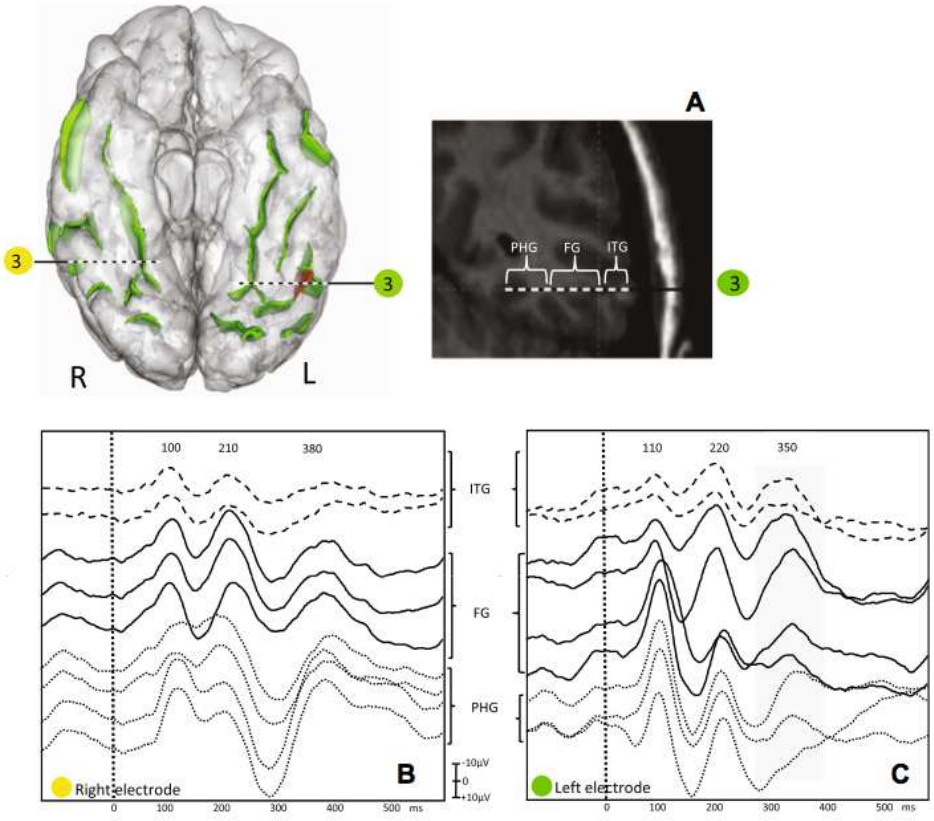


Figure 1.27: Electrode implantation and recordings with sEEG. (Reproduced from [72]).

Functional MRI can be used to produce volumetric activation maps showing which parts of the brain are involved in a particular mental process (cf. figure 1.28).

Oxygen is delivered to neurons via haemoglobin carried by red blood cells. Haemoglobin is diamagnetic when it is oxygenated while it is paramagnetic when deoxygenated. This difference in magnetic properties leads to small differences in the MR signal. Since blood oxygenation varies according to the levels of neural activity these differences can be used to detect brain activity. This type of MRI is known as blood oxygenation level dependent (BOLD) imaging.

One point to note is that the blood oxygenation increases following neural activation with a delay of a few seconds. Due to the indirect measure of neural activation, temporal resolution of fMRI is limited to the time scales of the measured hemodynamic processes. See [184] for a historical perspective on fMRI development.

Positron Emission Tomography (PET)

Positron emission tomography (PET) is a nuclear medicine imaging technique which produces a three-dimensional image of brain activations. The system detects pairs of gamma rays emitted indirectly by a positron-emitting radionuclide, a tracer, which is injected into the body on a biologically active molecule. Images of tracer concentration in 3D space within the brain are then reconstructed by computer analysis, as illustrated in figure 1.29. Without going into details, tracers used for brain PET scanning focus on the glucose consumption of the different brain regions. Like fMRI, it gives access to neural activity indirectly via the measurements of metabolic processes, but contrary to fMRI it requires the injection of an invasive radioactive tracer.

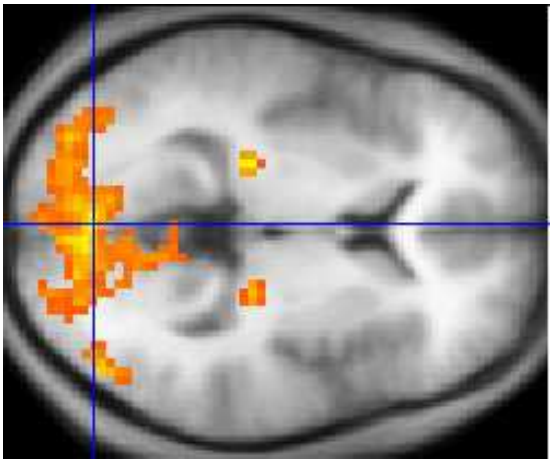


Figure 1.28: Sample fMRI activation map. The fMRI statistics (yellow) are overlaid on an average of the brain anatomies of several humans (source: wikipedia.org)

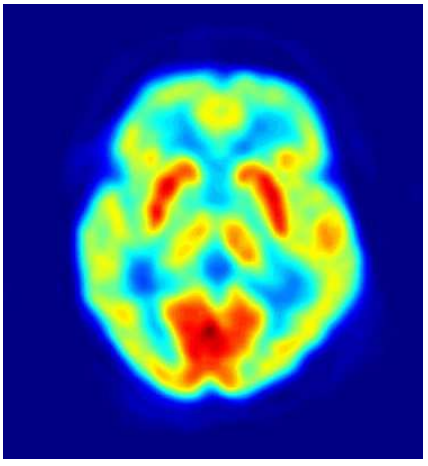


Figure 1.29: Sample PET activation map (source: wikipedia.org).

Single Photon Emission Computed Tomography (SPECT)

Single Photon Emission Computed Tomography (SPECT), is similar to PET in its use of radioactive tracer material and detection of gamma rays. In contrast to PET, however, the tracer used in SPECT emits gamma radiation that is measured directly, whereas PET tracer emits positrons which annihilate with electrons up to a few millimeters away, causing two gamma photons to be emitted in opposite directions. A PET scanner detects these emissions "coincident" in time, which provides more radiation event localization information and thus higher resolution images than SPECT. SPECT scans, however, are significantly less expensive than PET scans, in part because they are able to use longer-lived more easily-obtained radioisotopes than PET.

Optical imaging with near-infrared spectroscopy (nIRS)

Near-infrared spectroscopy (nIRS) uses near infrared light to measure the absorption of haemoglobin. It relies on the absorption spectrum of haemoglobin varying with its oxygenation status and as a consequence on the level of neural activity in a specific brain region. It has the interesting ability for measure both deoxygenated and oxygenated haemoglobin, which is of particular interest for understanding hemodynamic processes. nIRS is more convenient than fMRI with babies for which the skull is more transparent. It is also less noisy than fMRI systems and therefore more adapted to children. However, nIRS is restricted to superficial sources. See for example [19, 88] for more details.

1.3 Conclusion

As the electric activity of the neurons produces an electromagnetic field, and more importantly because, the organization of neural assemblies enables the summation of these fields, it is possible to detect and measure the brain activity outside of the head.

This offers the possibility to directly measure the neuronal activity in a non-invasive way. The high temporal resolution of M/EEG measurements makes them particularly interesting compared to other brain functional imaging modalities that are limited by the time scales of metabolic and hemodynamic processes. More than simply localizing the origin of the measured signal, M/EEG offer the possibility to investigate the dynamics of the cortical processing involved in different cognitive tasks.

In order to use M/EEG for brain functional imaging, some modeling and computation need to be done. Prior to any localization of activation, the way the neuronal currents and electromagnetic fields propagate within the different head tissues needs to be modeled. This aspect of the work, that consists in quantifying how the neuronal activations produce a signal on a given sensor involves physical considerations and is referred to as the *direct (or forward) problem*. This problem is the subject of the following chapter. The aspect that consists in localizing the activations based on the measured signal will be treated in chapters 3 and 4, and is called the *inverse problem*.

CHAPTER 2

THE FORWARD PROBLEM

In chapter 1, it was explained how neurons can induce electromagnetic fields that can be measured non-invasively outside of the head. The problem that consists in modeling the head in order to compute the electric potential or the magnetic field that should be produced by a given configuration of generators at the sensor level is called the forward problem. The solution of this problem is the first step in the M/EEG processing pipeline whose final objective is the localization of brain activations. The accuracy of the solution of the forward problem is fundamental in order to provide good localization results.

In the first part of this chapter, we review the equations and methods for solving the forward problem. We then detail the software contributions made in this thesis mainly in the OpenMEEG project that implements the symmetric BEM presented in this chapter. We also provide a list of open source software projects that can be used for solving the M/EEG forward problem. Finally, we provide some numerical evaluations that demonstrate that the precision obtained by OpenMEEG clearly improves over concurrent implementations.

Contents

2.1 THE PHYSICS OF EEG AND MEG

Notations

All vectors are denoted in bold characters. The vector indicating the position of a point r of $\mathbb{R}^3$ is denoted by $\mathbf{r}$. In the following, we use vector calculus notation, with the "nabla" operator ∇. For a real function $f(\mathbf{r})$, ∇f is the gradient of f. For a vector field $\mathbf{X}(\mathbf{r})$, $\nabla \cdot \mathbf{X}$ is the divergence of this field (a scalar) and $\nabla \times \mathbf{X}$ is the curl of this field (a vector).

2.1.1 Maxwell's equations

Maxwell's equations relate the electromagnetic field to the charge density and current density. We denote by $\mathbf{E}$ the electric field, $\mathbf{B}$ the magnetic field, ρ the charge density and $\mathbf{J}$ the current density. Maxwell's equations are a set of four partial differential equations:

$$\begin{cases} \nabla \cdot \mathbf{E} = \dfrac{\rho}{\epsilon} \\ \nabla \times \mathbf{E} = -\dfrac{\partial \mathbf{B}}{\partial t} \\ \nabla \cdot \mathbf{B} = 0 \\ \nabla \times \mathbf{B} = \mu \left(\mathbf{J} + \epsilon \dfrac{\partial \mathbf{E}}{\partial t} \right) \end{cases} \tag{2.1}$$

where ϵ is the electrical permittivity of the medium and μ is the magnetic permeability.

For human tissues, the magnetic permeability μ is the same as for vacuum $\mu = \mu_0$, whereas the relative electrical permittivity $\epsilon_r = \frac{\epsilon}{\epsilon_0}$ varies a lot depending on tissue and frequency. For instance, at a frequency of 100 Hz, ϵ_r is around 4×10^6 for gray matter, 5×10^5 for fat and 6×10^3 for compact bone [83].

2.1.2 Quasi-static approximation

As described in section 1.1, the post-synaptic potentials have a duration of about 10 ms. As a consequence, it is commonly accepted that the time frequencies of the brain electromagnetic field that can be observed outside the head can rarely exceed 100 Hz. For such low frequencies, the time derivatives in Maxwell's equations can be neglected, this is called the quasi-static approximation.

A justification of the quasi-static approximation can be found in [100]. Let us illustrate it with some orders of magnitude. We know that in a simple medium, the general solution of the electromagnetic wave equation can be written as a linear superposition of planar waves of different frequencies and polarizations. Let us just consider one planar wave for the sake of simplicity. Its equation is:

$$\mathbf{E}(\mathbf{r}, t) = \mathbf{E}_0 e^{i2\pi \mathbf{k} \cdot \mathbf{r}} e^{i2\pi f t} \quad , \tag{2.2}$$

where i is the imaginary unit, $\mathbf{E}_0$ is a real amplitude vector contained in the wave plane, $\mathbf{k}$ is a real spatial frequency vector normal to the wave plane ($\mathbf{E}_0 \cdot \mathbf{k} = 0$), and f is the temporal frequency. Let us consider a Maxwell's equation including a time derivative in a passive conductive non-magnetic medium with conductivity σ:

$$\nabla \times \mathbf{B} = \mu_0 \left(\sigma \mathbf{E} + \epsilon \frac{\partial \mathbf{E}}{\partial t} \right) \quad . \tag{2.3}$$

The current $\mathbf{J}$ is replaced by $\sigma\mathbf{E}$ in (2.1) following Ohm's law, $\mathbf{J} = \sigma\mathbf{E}$.

To neglect the time derivative in (2.3), it is necessary that $\|\epsilon\frac{\partial\mathbf{E}}{\partial t}\| \ll \|\sigma\mathbf{E}\|$. For the planar wave, this is equivalent to $\kappa = |2\pi f\frac{\epsilon}{\sigma}| \ll 1$. At a frequency of 100 Hz, the average permittivity of the head tissues is $\epsilon = 10^5\epsilon_0$ and the average conductivity is $\sigma = 0.3\ \Omega^{-1}\mathrm{m}^{-1}$. With these values, $\kappa = 1,8 \times 10^{-3}$, hence the term $\epsilon\frac{\partial\mathbf{E}}{\partial t}$ can be neglected.

More intuitively, we can just consider the spatial wavelength λ of our planar wave which is given by the relation $c = f\lambda$, where $c = \frac{1}{\sqrt{\mu\epsilon}}$ is the speed of the wave in the medium and f is the temporal frequency of the wave. With a frequency of 100 Hz, it gives us a wavelength of about 10^5 m. Thus, at the scale of a human head, we can neglect the oscillations of the wave, which gives $\nabla \times \mathbf{E} = 0$ instead of $\nabla \times \mathbf{E} = -\frac{\partial\mathbf{B}}{\partial t}$.

The practical consequences of the quasi-static approximation are twofold. First, the electric component is decoupled from the magnetic component, allowing the computation of the electric potential separately. Second, the delays of propagation of the signal from the neuronal sources to the M/EEG sensors can be neglected. We can, therefore, assume that M/EEG sensors measure at each instant the activity produced at the very same instant.

2.1.3 The electric potential equation

In the quasi-static approximation, we neglect all the time derivatives. As a consequence, the curl of the electric field $\mathbf{E}$ is zero, meaning that it derives from a scalar potential V:

$$\mathbf{E} = -\nabla V. \tag{2.4}$$

In a medium with current generators, the total current can be decomposed in two parts: a *primary current* flow $\mathbf{J}^{\mathrm{p}}$ related to the current generators and a volume current flow $\mathbf{J}^{\mathrm{v}}$ due to the electric field in the volume.

Using Ohm's Law, $\mathbf{J}^{\mathrm{v}} = \sigma\mathbf{E}$, we have:

$$\mathbf{J} = \mathbf{J}^{\mathrm{p}} + \mathbf{J}^{\mathrm{v}} = \mathbf{J}^{\mathrm{p}} + \sigma\mathbf{E} = \mathbf{J}^{\mathrm{p}} - \sigma\nabla V \ . \tag{2.5}$$

The volume currents, a.k.a., the ohmic currents, correspond to the displacements of charges due to the gradient of potential in the medium.

Neglecting the time derivative in Maxwell equations leads to:

$$\begin{aligned} & \nabla \times \mathbf{B} &&= \mu\mathbf{J} \\ \Rightarrow\ & \nabla \cdot (\nabla \times \mathbf{B}) &&= \nabla \cdot (\mu\mathbf{J}) \\ \Rightarrow\ & 0 &&= \nabla \cdot (\mu\mathbf{J}) \\ \Rightarrow\ & 0 &&= \nabla \cdot \mathbf{J} \end{aligned} \tag{2.6}$$

which finally leads to the potential equation:

$$\boxed{\nabla \cdot (\sigma\nabla V) = \nabla \cdot \mathbf{J}^{\mathrm{p}} \ .} \tag{2.7}$$

2.1.4 The magnetic field equation: the Biot-Savart law

Because $\nabla \cdot \mathbf{B} = 0$, there exists a vector field $\mathbf{A}$ such that:

$$\mathbf{B} = \nabla \times \mathbf{A} \ .$$

We use the classical gauge condition $\nabla \cdot \mathbf{A} = 0$ to avoid the indetermination caused by the definition of $\mathbf{A}$.

This leads to:

$$\nabla \times \mathbf{B} = \nabla \times \nabla \times \mathbf{A} = \nabla (\nabla \cdot \mathbf{A}) - \Delta \mathbf{A} = -\Delta \mathbf{A} \ .$$

Maxwell's equation $\nabla \times \mathbf{B} = \mu_0 \mathbf{J}$ becomes $\Delta \mathbf{A} = -\mu_0 \mathbf{J}$, which is a Poisson equation. If we impose $\mathbf{A}(|r| \to \infty) = 0$ (no magnetic field at infinity), it has a general solution in $\mathbb{R}^3$:

$$\mathbf{A}(\mathbf{r}) = \frac{\mu_0}{4\pi} \int_{\mathbb{R}^3} \frac{\mathbf{J}(\mathbf{r}')}{\|\mathbf{r} - \mathbf{r}'\|} d\mathbf{r}' \ .$$

Taking the curl, we obtain the Biot-Savart law:

$$\mathbf{B}(\mathbf{r}) = \frac{\mu_0}{4\pi} \int_{\mathbb{R}^3} \mathbf{J}(\mathbf{r}') \times \frac{(\mathbf{r} - \mathbf{r}')}{\|\mathbf{r} - \mathbf{r}'\|^3} d\mathbf{r}' \ .$$

Because the current can be written as $\mathbf{J} = \mathbf{J}^{\mathrm{p}} - \sigma \nabla V$, we can transform the Biot-Savart law into:

$$\boxed{\mathbf{B}(\mathbf{r}) = \mathbf{B}_0(\mathbf{r}) - \frac{\mu_0}{4\pi} \int_{\mathbb{R}^3} \sigma \nabla V(\mathbf{r}') \times \frac{\mathbf{r} - \mathbf{r}'}{\|\mathbf{r} - \mathbf{r}'\|^3} d\mathbf{r}' \ ,} \tag{2.8}$$

with

$$\mathbf{B}_0(\mathbf{r}) = \frac{\mu_0}{4\pi} \int_{\mathbb{R}^3} \mathbf{J}^{\mathrm{p}}(\mathbf{r}') \times \frac{(\mathbf{r} - \mathbf{r}')}{\|\mathbf{r} - \mathbf{r}'\|^3} d\mathbf{r}' \ .$$

With this formulation, $\mathbf{B}_0$ is often called the *primary* magnetic field while the second term is called the *secondary* magnetic field.

Note that the primary magnetic field does not depend on the medium, which corresponds to the head with M/EEG, and therefore it does not depend on the values of the conductivities.

2.2 UNBOUNDED HOMOGENEOUS MEDIUM

Let us consider an homogeneous volume conductor with a constant conductivity σ. The equation (2.7) becomes:

$$\Delta V = \frac{1}{\sigma} \nabla \cdot \mathbf{J}^{\mathrm{p}} \ ,$$

which is a Poisson equation of general solution:

$$V(\mathbf{r}) = \frac{1}{4\pi\sigma} \int_{\mathbb{R}^3} \frac{\nabla \cdot \mathbf{J}^{\mathrm{p}}(\mathbf{r}')}{\|\mathbf{r} - \mathbf{r}'\|} d\mathbf{r}' \ ,$$

with V vanishing at infinity. Applying the divergence theorem, it yields

$$\boxed{V(\mathbf{r}) = \frac{1}{4\pi\sigma} \int_{\mathbb{R}^3} \mathbf{J}^{\mathrm{p}}(\mathbf{r}') \cdot \frac{(\mathbf{r} - \mathbf{r}')}{\|\mathbf{r} - \mathbf{r}'\|^3} d\mathbf{r}' \ .} \tag{2.9}$$

For the magnetic field, we take σ out of the integral in (2.8) because it is constant:

$$\mathbf{B}(\mathbf{r}) = \mathbf{B}_0(\mathbf{r}) - \frac{\mu_0 \sigma}{4\pi} \int_{\mathbb{R}^3} \nabla V(\mathbf{r}') \times \frac{(\mathbf{r} - \mathbf{r}')}{\|\mathbf{r} - \mathbf{r}'\|^3} d\mathbf{r}' \ . \tag{2.10}$$

Using the identity

$$\nabla V(\mathbf{r}') \times \frac{(\mathbf{r} - \mathbf{r}')}{\|\mathbf{r} - \mathbf{r}'\|^3} = \nabla \times \left(\frac{\nabla V(\mathbf{r}')}{\|\mathbf{r} - \mathbf{r}'\|} \right) - \frac{\nabla \times (\nabla V(\mathbf{r}'))}{\|\mathbf{r} - \mathbf{r}'\|}$$

and the fact that the curl of a gradient is null, the integral on the right hand side of (2.10)

becomes

$$\int_{\mathbb{R}^3} \nabla \times \left(\frac{\nabla V(\mathbf{r}')}{\|\mathbf{r} - \mathbf{r}'\|} \right) d\mathbf{r}' \ .$$

Since V vanishes at infinity, using Stokes' theorem, this integral is null. We obtain finally that, in an infinite homogeneous medium, the magnetic field reduces to the primary field:

$$\boxed{\mathbf{B}(\mathbf{r}) = \mathbf{B}_0(\mathbf{r}) = \frac{\mu_0}{4\pi} \int_{\mathbb{R}^3} \mathbf{J}^{\mathrm{p}}(\mathbf{r}') \times \frac{(\mathbf{r}-\mathbf{r}')}{\|\mathbf{r}-\mathbf{r}'\|^3} d\mathbf{r}' \ .} \tag{2.11}$$

In this special case, the passive current $\sigma\mathbf{E} = -\sigma\nabla V$ does not contribute to the magnetic field.

2.2.1 Dipolar sources

If the primary current $\mathbf{J}^{\mathrm{p}}$ is reduced to a single current dipole at position $\mathbf{r}_0$ with moment $\mathbf{q}$, then $\mathbf{J}^{\mathrm{p}}(\mathbf{r}) = \mathbf{q}\ \delta_{\mathbf{r}_0}(\mathbf{r})$, where $\delta_{\mathbf{r}_0}$ is the Dirac distribution at $\mathbf{r}_0$. Using equations (2.9) and (2.11) with such a primary current, we obtain that the potential and magnetic field in an homogeneous space have very simple formulations:

$$V_{dip}(\mathbf{r}) = \frac{1}{4\pi\sigma} \mathbf{q} \cdot \frac{\mathbf{r} - \mathbf{r}_0}{\|\mathbf{r} - \mathbf{r}_0\|^3} \tag{2.12}$$

$$\mathbf{B}_{dip}(\mathbf{r}) = \frac{\mu_0}{4\pi} \mathbf{q} \times \frac{\mathbf{r} - \mathbf{r}_0}{\|\mathbf{r} - \mathbf{r}_0\|^3} \tag{2.13}$$

Such formulas can also be obtained from a Taylor expansion of the function:

$$\Phi_{\mathbf{r}}(\mathbf{r}') = \frac{\mathbf{r} - \mathbf{r}'}{\|\mathbf{r} - \mathbf{r}'\|^3}$$

Let us assume that the primary current sources lie in small volume $\delta\Omega$ (cf. figure 2.1). The Taylor expansion of $\Phi_{\mathbf{r}}$ at $\mathbf{r}' = \mathbf{r}_0$, where $\mathbf{r}_0$ is the centroid of the small volume $\delta\Omega$, gives:

$$\Phi_{\mathbf{r}}(\mathbf{r}') = \Phi_{\mathbf{r}}(\mathbf{r}_0) + \nabla_{\mathbf{r}_0}\Phi_{\mathbf{r}}(\mathbf{r}_0 - \mathbf{r}') + o(\|\mathbf{r}' - \mathbf{r}_0\|)$$

where $\nabla_{\mathbf{r}_0}\Phi_{\mathbf{r}}$ is the gradient of Φ_r taken at position $\mathbf{r}_0$.

By approximating $\Phi_{\mathbf{r}}(\mathbf{r}')$ by $\Phi_{\mathbf{r}}(\mathbf{r}_0)$ the equations (2.9) and (2.11) write:

$$\begin{aligned} V(\mathbf{r}) &= \frac{1}{4\pi\sigma} \int_{\delta\Omega} \mathbf{J}^{\mathrm{p}}(\mathbf{r}') d\mathbf{r}' \cdot \frac{\mathbf{r} - \mathbf{r}_0}{\|\mathbf{r} - \mathbf{r}_0\|^3} \\ \mathbf{B}(\mathbf{r}) &= \frac{\mu_0}{4\pi} \int_{\delta\Omega} \mathbf{J}^{\mathrm{p}}(\mathbf{r}') d\mathbf{r}' \times \frac{\mathbf{r} - \mathbf{r}_0}{\|\mathbf{r} - \mathbf{r}_0\|^3} \ . \end{aligned} \tag{2.14}$$

We obtain the formulas giving the potential and magnetic field of a dipolar source whose moment is given by:

$$\mathbf{q} = \int_{\delta\Omega} \mathbf{J}^{\mathrm{p}}(\mathbf{r}') d\mathbf{r}' \ . \tag{2.15}$$

As illustrated in figure 2.1, this approximation is justified if $\|\mathbf{r}' - \mathbf{r}_0\| \ll \|\mathbf{r}' - \mathbf{r}\|$ in $\delta\Omega$. This implies that the primary currents produced by a region $\delta\Omega$ sufficiently small compared to the distance to the observation point $\mathbf{r}$ can be correctly modeled by an equivalent current dipole (ECD).

2.2.2 Multipolar sources

If the Taylor approximation at order 0 used to justify the dipolar source model does not hold, it is necessary to consider more terms in the Taylor expansion. This leads to the multipolar

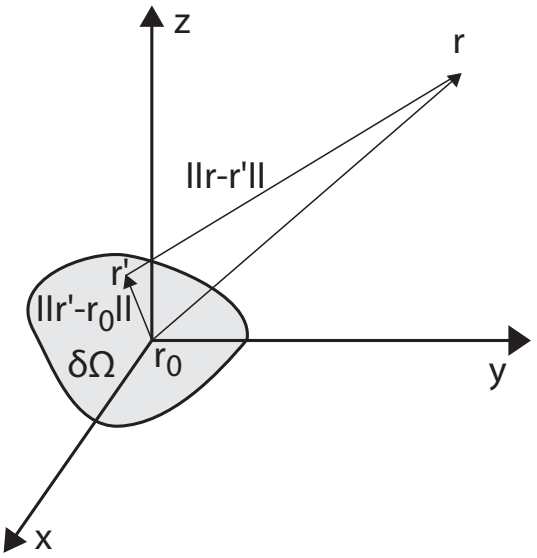


Figure 2.1: Measurement at point $\mathbf{r}$ of the electromagnetic fields produced by a current distribution in a region $\delta\Omega$ when $\|\mathbf{r}' - \mathbf{r}_0\| \ll \|\mathbf{r}' - \mathbf{r}\|$.

source models.

Using one more term we get:

$$\begin{aligned} V_{mult}(\mathbf{r}) &= \frac{1}{4\pi\sigma}\left(\int_{\delta\Omega}\mathbf{J}^{\mathrm{p}}(\mathbf{r}')d\mathbf{r}'\cdot\frac{\mathbf{r}-\mathbf{r}_0}{\|\mathbf{r}-\mathbf{r}_0\|^3}+\int_{\delta\Omega}\nabla_{\mathbf{r}_0}\Phi_{\mathbf{r}}(\mathbf{r}_0-\mathbf{r}')\cdot\mathbf{J}^{\mathrm{p}}(\mathbf{r}')d\mathbf{r}'\right) \\ &= V_{dip}(\mathbf{r})+V_{quad}(\mathbf{r}) \end{aligned} \tag{2.16}$$

where V_{quad} stands for the *quadrupolar* term.

Borrowing the notation ":" for *tensor contraction* from [118], we rewrite the quadrupolar term $V_{quad}(\mathbf{r})$ as:

$$V_{quad}(\mathbf{r}) = \frac{1}{4\pi\sigma}\nabla_{\mathbf{r}_0}\Phi_{\mathbf{r}} : \int_{\delta\Omega}(\mathbf{r}_0-\mathbf{r}')\mathbf{J}^{\mathrm{p}}(\mathbf{r}')d\mathbf{r}'$$

where the term $\int_{\delta\Omega}(\mathbf{r}_0-\mathbf{r}')\mathbf{J}^{\mathrm{p}}(\mathbf{r}')d\mathbf{r}'$, denoted by $\mathbf{Q}_{\boldsymbol{quad}}$, is called the quadrupolar moment. This term is a 3×3 tensor.

Similarly, for the magnetic field we get:

$$\mathbf{B}_{mult}(\mathbf{r}) = \mathbf{B}_{dip}(\mathbf{r}) + \mathbf{B}_{quad}(\mathbf{r})$$

where

$$\mathbf{B}_{quad}(\mathbf{r}) = \frac{\mu_0}{4\pi}\nabla_{\mathbf{r}_0}\Phi_{\mathbf{r}} : \mathbf{Q}_{\boldsymbol{quad}}$$

When considering more terms in the expansion, we add correcting terms in the expression of the electric and magnetic fields.

2.3 THE SPHERICALLY SYMMETRIC HEAD MODEL

Obviously, the human head is not an infinite homogeneous conductor. First of all, it is a bounded conductor and no electric current can flow outside the head (except at the neck). Secondly, the electrical conductivity σ of the head is not constant: for instance, the skull is

between 20 and 100 times less conductive than other head tissues. This must be taken into account to get more accurate approximations of the potential and magnetic field generated by the brain electrical activity. A first step towards head modeling is to consider the head as a set of nested concentric spheres. Each volume enclosed between two spheres is supposed to represent a different tissue with a constant isotropic conductivity. Figure 2.2 shows a sphere model with three spheres. Without respect to proportions, it could represent the brain, the skull and the scalp of a human head. This simple geometry allows one to find an analytic solution for the electric potential generated by a dipole, like for the infinite homogeneous medium (2.12).

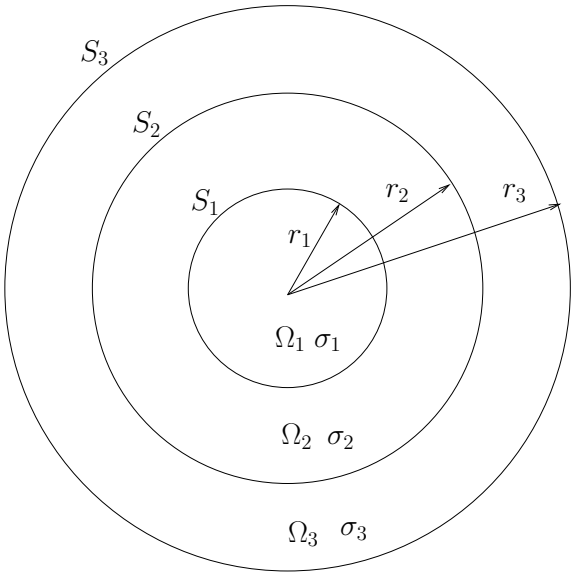


Figure 2.2: A spherical model with three layers.

2.3.1 Electric potential generated by a dipole

The key point is to take advantage of the spherical symmetry of the geometry. First, we use spherical coordinates (r, θ, ϕ) instead of Cartesian coordinates, and second, we expand the electric potential in spherical harmonics $Y_l^m(\theta, \phi)$. The spherical harmonics have the following form:

$$Y_l^m(\theta, \phi) = N_l^m P_l^m(\cos\theta) e^{im\phi}, \;\; l \in \mathbb{N}, m \in \mathbb{Z}, \;\; |m| < l \;\; ,$$

where N_l^m is a normalization coefficient and P_l^m is an associated Legendre function. In our case, this is of particular interest because the general solution of Laplace's equation $\Delta f = 0$ in spherical coordinates can be written as a linear combination of spherical harmonics

$$f(r, \theta, \phi) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} (A_{lm} r^{-1-l} + B_{lm} r^l) Y_l^m(\theta, \phi) \;\; .$$

If f is a real function, it simplifies to

$$f(r, \theta, \phi) = \sum_{l=0}^{\infty} \sum_{m=0}^{l} (A_{lm} r^{-1-l} + B_{lm} r^l) P_l^m(\cos\theta) \cos(m\phi) \;\; .$$

Now let us consider a current dipole inside the spherical model at location $\mathbf{r}_0$ with $\mathbf{q}$ its moment. We use the notation of figure 2.2, with indices increasing from the innermost sphere

to the outermost one. In all subregions Ω_k where the dipole is not located, the potential equation states that $\nabla \cdot (\sigma_k \nabla V) = \sigma_k \Delta V = 0$ because σ_k is constant. As a consequence, the restriction V_k of V in each domain Ω_k is harmonic and can be decomposed on the spherical harmonic basis:

$$\boxed{V_k(r,\phi,\theta) = \sum_{l=0}^{\infty} \sum_{m=0}^{l} (A_{lm}^k r^{-1-l} + B_{lm}^k r^l) P_l^m(\cos\theta)\cos(m\phi) \ .} \tag{2.17}$$

In the domain Ω_{k^*} where the dipole is located, the potential V satisfies $\sigma_{k^*}\Delta V = \nabla \cdot \mathbf{J}^{\mathrm{p}}$, with $\mathbf{J}^{\mathrm{p}} = \mathbf{q}\ \delta_{\mathbf{r}_0}$. So we can decompose the potential in $V = v + u$, where v is the potential generated by the dipole in an infinite homogeneous domain of conductivity σ_{k^*}, and u is an harmonic function. The function v is defined as

$$v(\mathbf{r}) = \frac{1}{4\pi\sigma_{k^*}} \mathbf{q} \cdot \frac{(\mathbf{r} - \mathbf{r}_0)}{\|\mathbf{r} - \mathbf{r}_0\|^3} \ .$$

This function can be decomposed in the spherical harmonic basis

$$v(r,\phi,\theta) = \begin{cases} \sum_{l=0}^{\infty} \sum_{m=0}^{l} q_{lm}^{inf} r^l P_l^m(\cos\theta)\cos(m\phi) \ , & r < r_0 \\ \sum_{l=0}^{\infty} \sum_{m=0}^{l} q_{lm}^{sup} r^{-1-l} P_l^m(\cos\theta)\cos(m\phi) \ , & r > r_0 \end{cases}$$

So if we denote $A_{lm}^{k^*}$ and $B_{lm}^{k^*}$ the coefficients of the decomposition of u, we have a decomposition of V_{k^*} in the spherical harmonic basis

$$\boxed{V_{k^*}(r,\phi,\theta) = \begin{cases} \sum_{l=0}^{\infty} \sum_{m=0}^{l} (A_{lm}^{k^*} r^{-1-l} + (q_{lm}^{inf} + B_{lm}^{k^*}) r^l) P_l^m(\cos\theta)\cos(m\phi) \ , & r < r_0 \\ \sum_{l=0}^{\infty} \sum_{m=0}^{l} ((A_{lm}^{k^*} + q_{lm}^{sup}) r^{-1-l} + B_{lm}^{k^*} r^l) P_l^m(\cos\theta)\cos(m\phi) \ , & r > r_0 \end{cases}} \tag{2.18}$$

Finally, to fully determine the potential in the whole domain Ω, one needs to fix the value of the coefficients A_{lm}^k and B_{lm}^k. This is done by considering the boundary conditions at each surface S_k. The electric potential and the current density must be continuous through the interfaces:

$$\begin{cases} V_k(r_k,\phi,\theta) = V_{k+1}(r_k,\phi,\theta) \\ \sigma_k \frac{\partial V_k}{\partial r}(r_k,\phi,\theta) = \sigma_{k+1} \frac{\partial V_{k+1}}{\partial r}(r_k,\phi,\theta) \end{cases} \tag{2.19}$$

From (2.17), (2.18) and (2.19), a linear system can be built for the A_{lm}^k and B_{lm}^k, which leads to the determination of these coefficients. Because of the infinite series, in practical situations one has to choose at which order the series have to be truncated. For high orders, the solution is more accurate but the computation is more expensive. Several approaches have been proposed for efficient computation of the electric potential in multilayer spheres [18, 58, 238].

2.3.2 The magnetic field

In the case of the magnetic field, there is no need to use an infinite series based on a decomposition in spherical harmonics. Indeed, in a spherical geometry, the magnetic field has a simple closed-form.

2.3.2.1 The radial component of the magnetic field

We again consider a spherical geometry as described in figure 2.2. With spherical coordinates, the radial component of the magnetic field is

$$\mathbf{B}_r(\mathbf{r}) = \mathbf{B}(\mathbf{r}) \cdot \mathbf{e}_r = \mathbf{B}(\mathbf{r}) \cdot \frac{\mathbf{r}}{r} \ ,$$

and the outward normal at each surface S_k is

$$\mathbf{n}(\mathbf{r}') = \frac{\mathbf{r}'}{r'} \ .$$

The spherical geometry in figure 2.2 is a special case of geometry with a piecewise constant conductivity, so that the magnetic field can be expressed using the formula (2.28) described in section 2.4.3. If we compute the radial component of the magnetic field, the following scalar triple product appears in the surface integrals:

$$\frac{\mathbf{r} - \mathbf{r}'}{\|\mathbf{r} - \mathbf{r}'\|^3} \times \frac{\mathbf{r}'}{r'} \cdot \frac{\mathbf{r}}{r} \ .$$

This quantity is zero because $\mathbf{r}$, $\mathbf{r}'$ and $(\mathbf{r} - \mathbf{r}')$ are in a same plane. As a consequence, in a spherical geometry, the radial component of the magnetic field is equal to the radial component of the primary field:

$$\mathbf{B}_r(\mathbf{r}) = \mathbf{B}_0(\mathbf{r}) \cdot \mathbf{e}_r = \frac{\mu_0}{4\pi} \int_\Omega \mathbf{J}^{\mathrm{p}} \times \frac{(\mathbf{r} - \mathbf{r}')}{\|\mathbf{r} - \mathbf{r}'\|^3} \cdot \mathbf{e}_r d\mathbf{r}' \ . \tag{2.20}$$

2.3.2.2 Total magnetic field generated by a dipole

We assume that $\mathbf{J}^{\mathrm{p}}$ is a dipole at position $\mathbf{r}_0$ with a moment $\mathbf{q}$. Outside the domain Ω, there is no current, the Maxwell's equations in the quasi-static approximation state that $\nabla \times \mathbf{B} = 0$. As a consequence, outside Ω, $\mathbf{B}$ derives from a scalar potential U:

$$\mathbf{B} = -\nabla U \ ,$$

with U vanishing at infinity. For $\mathbf{r}$ outside Ω, we can then write the following line integral:

$$\begin{aligned} U(\mathbf{r}) &= -\int_0^\infty \nabla U(\mathbf{r} + t\mathbf{e}_r) \cdot \mathbf{e}_r dt \\ &= \int_0^\infty \mathbf{B}(\mathbf{r} + t\mathbf{e}_r) \cdot \mathbf{e}_r dt \\ &= \int_0^\infty \mathbf{B}_r(\mathbf{r} + t\mathbf{e}_r) \cdot \mathbf{e}_r dt \end{aligned}$$

Using the expression of the radial magnetic field in (2.20) and the dipolar approximation from (2.15), we obtain:

$$U(\mathbf{r}) = \frac{\mu_0}{4\pi} \mathbf{q} \times (\mathbf{r} - \mathbf{r}_0) \cdot \mathbf{e}_r \int_0^\infty \frac{1}{\|\mathbf{r} + t\mathbf{e}_r - \mathbf{r}_0\|^3} dt \ .$$

The computation of the integral in the right hand side leads to

$$U(\mathbf{r}) = -\frac{\mu_0}{4\pi} \frac{\mathbf{q} \times \mathbf{r}_0 \cdot \mathbf{r}}{F} \ ,$$

where $F = a(ra + r^2 - \mathbf{r}_0 \cdot \mathbf{r})$ with $\mathbf{a} = \mathbf{r} - \mathbf{r}_0$ and $a = \|\mathbf{a}\|$. Taking the gradient results in the following formulation for the total magnetic field:

$$\boxed{\mathbf{B}(\mathbf{r}) = \frac{\mu_0}{4\pi F^2}(F\mathbf{q} \times \mathbf{r}_0 - \mathbf{q} \times \mathbf{r}_0 \cdot \mathbf{r}\nabla F) \ .} \tag{2.21}$$

This formula for the total magnetic field generated by a dipole in a spherical geometry was found by Sarvas [192]. Interestingly, although it is different from the formula in an infinite homogeneous medium, it is also independent of the conductivity σ of the domain Ω.

2.3.3 Magnetic field generated by a multipole

Considering multipolar expansions, Jerbi *et al.* showed in [118], that a similar expression as (2.21) can be derived for the magnetic field when using spherical geometries.

Let $\mathbf{X_r}$ be the cross product tensor, defined by $\mathbf{r} \times \mathbf{x} = \mathbf{X_r} \cdot \mathbf{x}$. Using this notation, the Sarvas formula can be rewritten for a current dipole:

$$\mathbf{B}_{\boldsymbol{dip}}(\mathbf{r}) = -\frac{\mu_0}{4\pi F^2}(F\mathbf{X}_{\mathbf{r_0}} + \nabla F(\mathbf{r}_0 \times \mathbf{r})) \cdot \mathbf{q} \ .$$

Borrowing again the notation ":" from [118], the magnetic field produced by a current multipole can be obtained from:

$$\mathbf{B}_{\boldsymbol{mult}}(\mathbf{r}) = -\frac{\mu_0}{4\pi F^2}(F\mathbf{X}_{\mathbf{r_0}} + \nabla F(\mathbf{r}_0 \times \mathbf{r})) \cdot \mathbf{q} - \nabla_{\mathbf{r}_0}\frac{\mu_0}{4\pi F^2}(F\mathbf{X}_{\mathbf{r_0}} + \nabla F(\mathbf{r}_0 \times \mathbf{r})) : \mathbf{Q}_{quad} \tag{2.22}$$

The term $\mathbf{q}$ has 3 coefficients and $\mathbf{Q}_{quad}$ has 9 coefficients. Yet, an analysis of this first order multipole model [117] has shown that the forward field produced is only of rank 7. A first order current multipole can therefore be modeled with 7 parameters instead of 12.

2.3.4 Limits of spherical models

The analytical or semi-analytical formulas of the electromagnetic field can be extended to non concentric spheres [149], or to ellipsoidal geometries [54]. However, for EEG, several studies have shown that such simplified models can not produce satisfactory results [33, 47, 109]. For MEG data, the head modeling is not as crucial. With a spherical head model, the total magnetic field does not depend on the conductivities and it is observed with more complex head geometries that this limited influence of the conductivities is maintained. This explains why spherical models are very popular in MEG. With EEG, it is necessary to considerer realistic head models with less constrained geometrical properties.

2.4 Realistic head models

To improve the accuracy of the forward calculation one needs to consider more realistic head models. The geometry of such improved head models can be obtained from other anatomical imaging modalities: computed tomography (CT) and structural magnetic resonance imaging (sMRI). Structural MRI is here opposed to functional MRI (fMRI) used for brain functional imaging with MRI (cf. section 1.2.3).

MRI consists in applying a strong external magnetic field to a volume, which consequently aligns all the magnetic moments of nucleus like hydrogen with a fixed direction within this volume. After stopping this magnetic field, MR machines are able to measure the relaxation

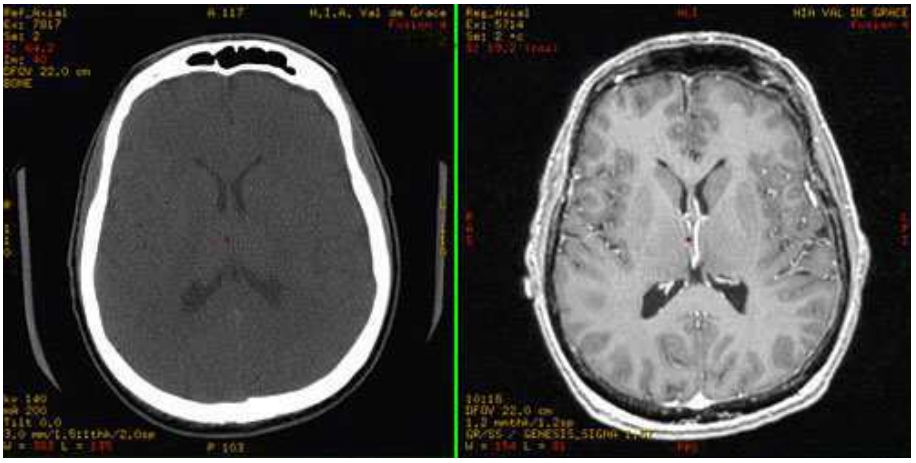


Figure 2.3: On the left, a slice of a CT image. On the right, the same slice obtained with T1 MRI. We observe that CT offers a clear view of the skull while MRI provides a good contrast in soft tissues. (Source gehealthcare.com)

time in voxels located on a 3D grid within the volume. The relaxation time is the time it takes for the hydrogen to return to equilibrium. This relaxation time depends of the physical properties of each tissue offering the possibility to get 3D images with a high contrast especially for soft tissues (white matter, gray matter, fat, muscle). MRI is however not very adapted to the imaging of bones due to the reduced presence of hydrogen in such structures. Computed tomography (CT) imaging modality is more appropriate for bones. However, based on X-ray, this modality exposes the patients to the hazards of ionizing radiation and is therefore not commonly used for M/EEG studies. Hence, precise models of the head tissues are often built from structural MR images, but the skull is most of the time not clearly visible and the skull models obtained from MRI are less accurate. CT and MRI images of the same subject are shown in figure 2.3.

Volumetric anatomical data reveal the geometrical complexity of the head structures. We will now present existing approaches and approximations to take this information into account for computing precise forward models. The different approaches that exist to compute numerical solutions for the forward problem in M/EEG are Finite Difference Methods (FDM), Finite Element Methods (FEM) and the Boundary Element Method (BEM).

2.4.1 The Finite Difference Method (FDM)

Finite difference methods provide numerical solutions to differential equations by approximating derivatives with finite differences, *i.e.,* approximative equivalent difference quotients.

For instance, for a function f in 1D, the first order derivative is given by the limit:

$$f'(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h} \ ,$$

thus, for a small value of h, the derivative can be approximated by:

$$f'(x) \simeq \frac{f(x+h) - f(x)}{h} \ .$$

For equation (2.7), we need to approximate the differential operator $\nabla \cdot (\sigma \nabla V)$. In 3D, a point has 6 neighbors, located at a distance of $+h$ and $-h$ in each direction. This approxima-

tion leads to:

$$(\nabla \cdot \sigma \nabla V)(\mathbf{r}_0) \simeq \frac{1}{h^2}\left(\alpha_0 V(\mathbf{r}_0) - \sum_{i=1}^{6} \alpha_i V(\mathbf{r}_i)\right) \quad , \tag{2.23}$$

where the constants α_0 and α_i depend on the conductivities at the points $\mathbf{r}_0$ and $\mathbf{r}_i$. Please note that this scheme corresponds exactly to Kirchhoff's law for the balance of currents, assuming that the points form a network of resistors. Generally, the head volume is discretized using a cubic grid with a regular spacing h, therefore the same scheme (2.23) can be used at every point of the grid by computing differences between closest neighbors.

For the source, we need to approximate the divergence operator $\nabla \cdot \mathbf{J}^{\mathrm{p}}$. The primary currents are defined over the edges between the grid points. For example, a dipole can be represented as a small current flowing over the edge linking two points $\mathbf{r}^+$ and $\mathbf{r}^-$, so that the divergence is reduced to the source and sink of current, *i.e.,* $\nabla \cdot \mathbf{J}^{\mathrm{p}} = I\delta_{\mathbf{r}^+} - I\delta_{\mathbf{r}^-}$, where I is the amplitude of the current. Denoting by $[J_i]$ the values of the primary currents between the neighboring grid points, the term $\nabla \cdot \mathbf{J}^{\mathrm{p}}$ can be written in matrix form as $\mathbf{B}\,[J_i]$. By denoting $[V_i]$ the values of the potential at grid points and plugging this expression into Kirchhoff's law, we get that the potential $[V_i]$ is solution of the linear problem:

$$\mathbf{A}\,[V_i] = \mathbf{B}\,[J_i]$$

The matrices involved are typically very large since the whole head domain has to be discretized. However the matrix $\mathbf{A}$ that needs to be inverted is highly sparse, because it has at most six off-diagonal elements per line, which implies that iterative methods are efficient.

The main drawback of the FDM method for M/EEG forward modeling is that, due to the cubic grid, the complex interfaces between brain structures and thin layers cannot be precisely modeled. Indeed, with a cubic grid, the interfaces have to follow the grid points which leads to a "staircase" effect.

2.4.2 The Finite Element Method (FEM)

The FEM can work on an unstructured grid like a triangulated surface when computed in 2D or on a tetrahedrized volume in 3D. Figure 2.4 is an example of a tetrahedral mesh of the head. The problem is first reformulated in its variational form, also called *weak* form. Then, an approximate solution of the problem in its weak formulation is found by looking for a solution in a finite dimensional vector space. The second step requires to properly choose the finite dimensional space in order to guarantee the quality of the approximation. Once discretized, the problem leads to a linear system that is solved numerically. Contrary to the FDM, Finite Element Methods (FEM) do not suffer from the staircase effect.

To illustrate this, let us consider the equation (2.7) for the potential. We consider the domain Ω describing the head and its boundary denoted $\partial\Omega$. On the boundary, there is no electric current flowing outside, so the differential equation with its boundary condition is

$$\begin{cases} \nabla \cdot (\sigma \nabla V) = \nabla \cdot \mathbf{J}^{\mathrm{p}} \text{ in } \Omega \\ \sigma \nabla V \cdot \mathbf{n} = 0 \text{ on } \partial\Omega \end{cases} \tag{2.24}$$

An important step of the FEM is to transform the differential equation (2.24) in its variational formulation. We assume that V lives in a certain Hilbert space E of regular functions.

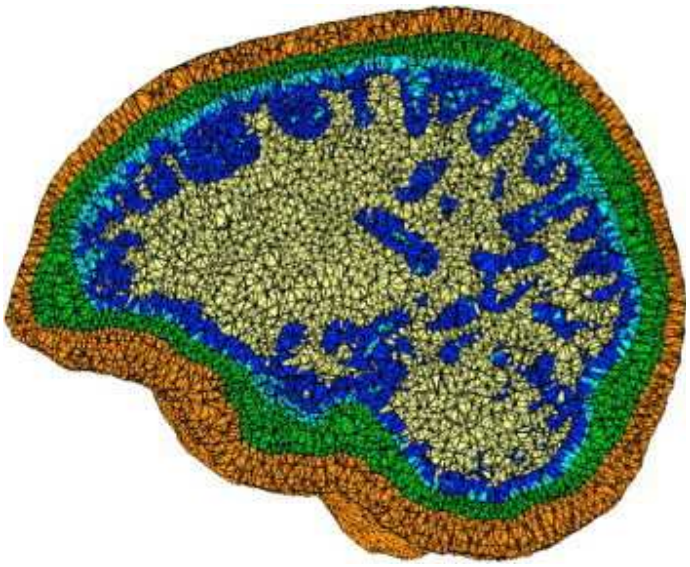


Figure 2.4: A tetrahedral mesh of the head. The different domains are shown with different colors, from inside to outside: white matter, gray matter, CSF, skull, scalp (Adapted from [229]).

If V is solution of (2.24), then for any function ϕ in E:

$$\begin{aligned}\int_\Omega \nabla \cdot \mathbf{J}^\mathrm{p} \phi &= \int_\Omega \nabla \cdot (\sigma \nabla V)\, \phi \\ &= \int_{\partial\Omega} \phi \sigma \nabla V \cdot \mathbf{n} - \int_\Omega \sigma \nabla V \cdot \nabla \phi \\ &= -\int_\Omega \sigma \nabla V \cdot \nabla \phi\end{aligned}$$

The last equality is of the form $a(V,\phi) = f(\phi)$ where a is a bilinear functional and f is linear. This provides the weak formulation of (2.24):

$$\forall \phi \in E, \quad a(V,\phi) = f(\phi) \ . \tag{2.25}$$

The second step consists in solving equation (2.25) in a finite dimensional subspace E_h of E. The parameter h refers to the precision of the approximation. Let $(\phi_i)_{i=1\ldots n}$ be a basis of E_h. A solution V in E_h can be written $V = \sum_{i=1}^n V_i \phi_i$.

The variational formulation in E_h becomes:

$$\forall j \in [1,\ldots,n], \quad \sum_{i=1}^n V_i a_h(\phi_i,\phi_j) = f_h(\phi_j) \ . \tag{2.26}$$

The right hand side is given by:

$$\begin{aligned} f_h(\phi_j) &= \textstyle\int_{\Omega_h} \nabla \cdot \mathbf{J}^\mathrm{p} \phi_j \\ &= \textstyle\int_{\partial\Omega_h} \phi_j \mathbf{J}^\mathrm{p} \cdot \mathbf{n} - \int_{\Omega_h} \mathbf{J}^\mathrm{p} \cdot \nabla \phi_j \\ &= \textstyle -\int_{\Omega_h} \mathbf{J}^\mathrm{p} \cdot \nabla \phi_j \ , \end{aligned}$$

assuming that there are no sources on the boundary $\partial\Omega_h$ and that partial integration is possible. For instance, for a dipole, $\mathbf{J}^\mathrm{p} = \mathbf{q}\delta_{\mathbf{r}_0}$, where $\delta_{\mathbf{r}_0}$ is the Dirac distribution at $\mathbf{r}_0$, the equality

$\int_{\Omega_h} \nabla \cdot \mathbf{J}^{\mathrm{p}} \phi_j = -\int_{\Omega_h} \mathbf{J}^{\mathrm{p}} \cdot \nabla \phi_j$ leads to:

$$f_h(\phi_j) = \mathbf{q} \cdot \nabla \phi_j(\mathbf{r}_0) \ .$$

Like for the FDM, this leads to a linear system $\mathbf{A}\,[V_i] = \mathbf{b}$ which completely determines the values V_i. An approximate solution for the potential can then be obtained by solving this linear system which can however be huge. The computation of the magnetic field is not detailed here, but it can be obtained from the computed electric potential.

The success of the FEM comes from the idea that by using basis functions ϕ_i with a local support, the matrix $\mathbf{A}$ can be very sparse. Iterative methods, like conjugate gradient methods, can then perform well for solving this linear system.

Typical choices for the functions ϕ_i include the P0 and P1 elements. For instance, the the subspace of the piecewise constant functions corresponds to a P0 discretization. With P0 elements the ϕ_i are indexed by the tetrahedron and are constant on the ith tetrahedron. With P1 elements, the functions ϕ_i are indexed by the nodes of the tetrahedrization. Each function ϕ_i takes the value 1 at node i, 0 at every other node, and is affine on the tetrahedra adjacent to node i. With such compactly supported elements, we observe that $a_h(\phi_i, \phi_j) = 0$ for two non neighboring elements, which guarantees the sparsity of $\mathbf{A}$. As shown by (2.8), once the potential V is computed, the magnetic field can be obtained by numerical integration.

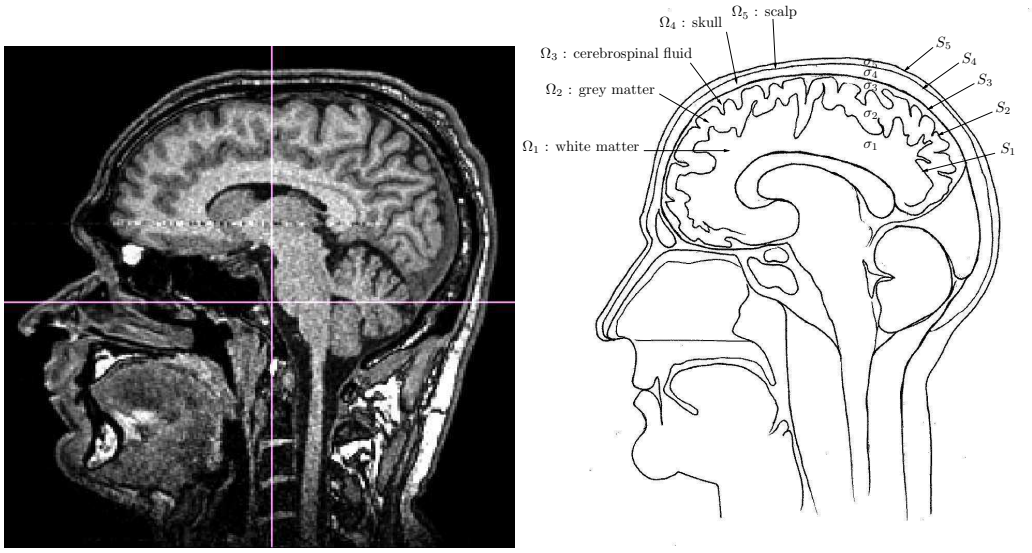
For the sake of clarity in the current presentation, the potential V is decomposed over a set of basis functions $(\phi_i)_i$, and it is these same functions that are used as test functions, a.k.a., evaluation functions. The approximation of (2.25) leads to (2.26) where only appear the dot products $a_h(\phi_i, \phi_j)$. However, it is also possible to use different basis of functions for the approximation of V and for the evaluation. If we consider different evaluation functions $(\psi_j)_j$, the discretization leads to the computation of $a_h(\phi_i, \psi_j)$. A particular choice for these functions leads to what is called collocation methods. With such methods, the test functions $(\psi_j)_j$ are Dirac functions and the computation of $a_h(\phi_i, \psi_j)$ leads to simple function evaluation. A numerical method where P1 elements are used for the approximation and Dirac functions are used for the evaluation is called "linear collocation" [155]. A collocation method with P0 elements for the discretization of V is called "constant collocation". Methods with general test functions are referred to as "Galerkin methods". For example, we call "linear Galerkin" (resp. "constant Galerkin") a method where both the approximation and the test functions are P1 (resp. P0) elements.

2.4.3 The Boundary Element Method (BEM)

From a structural MRI of the subject, it is possible to extract the different structures of the head (white matter, gray matter, *etc*.). In a first approximation, we can consider that the conductivity within each of these structures is constant. The boundary element method (BEM) is a numerical method for solving linear partial differential equations which have been transformed into integral equations defined over the boundaries of the different domains (white matter, gray matter, *etc*.). In order to achieve such a reformulation of the problem, one needs to assume homogeneous conductivities of each domain.

The piecewise constant approximation

Figure 2.5 shows the kind of geometry that can be extracted from the MRI of a subject's head. In practice, each subregion corresponds to a certain type of head tissue which is supposed to be sufficiently homogeneous to have a constant conductivity. The conductivity of the head is only discontinuous at the interfaces between tissues.



(a) Sagittal cross section of an anatomical MR volume image.

(b) The domains extracted from an MRI.

Figure 2.5: Example of piecewise constant head model.

With this approximation, we can model the head as a domain Ω composed of several subregions Ω_k separated by surfaces S_k, each with a constant conductivity σ_k, and with $\sigma = 0$ outside Ω. With the piecewise constant approximation, equations (2.7) and (2.8) can be transformed into integral equations:

$$\frac{\sigma_k + \sigma_{k+1}}{2} V(\mathbf{r}) = V_0(\mathbf{r}) - \frac{1}{4\pi} \sum_l (\sigma_l - \sigma_{l+1}) \int_{S_l} V(\mathbf{r}') \frac{\mathbf{r}-\mathbf{r}'}{\|\mathbf{r}-\mathbf{r}'\|^3} \cdot \mathbf{n}_l(\mathbf{r}') ds' \ , \tag{2.27}$$

$$\mathbf{B}(\mathbf{r}) = \mathbf{B}_0(\mathbf{r}) - \frac{\mu_0}{4\pi} \sum_l (\sigma_l - \sigma_{l+1}) \int_{S_l} V(\mathbf{r}') \frac{(\mathbf{r}-\mathbf{r}')}{\|\mathbf{r}-\mathbf{r}'\|^3} \times \mathbf{n}_l(\mathbf{r}') ds' \ , \tag{2.28}$$

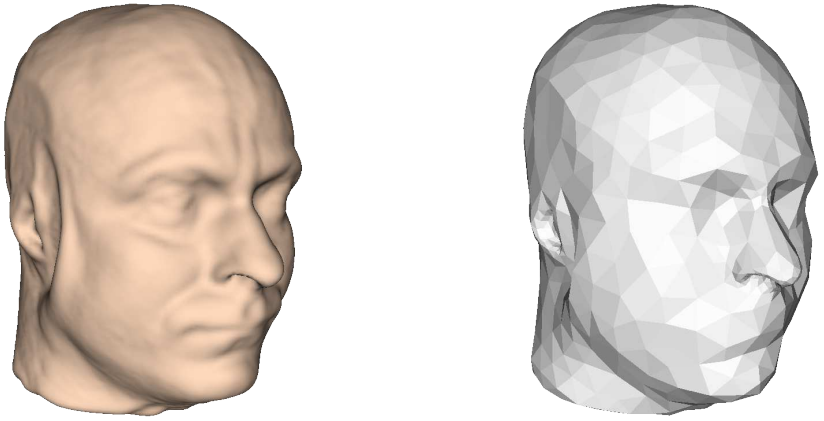
where $\mathbf{r} \in S_k$, V_0 and $\mathbf{B}_0$ are the electric potential and magnetic field generated by the primary current distribution $\mathbf{J}^{\mathrm{p}}$ in a homogeneous domain. These formulas are obtained from Geselowitz [86, 87]. Details on how to derive these equations are given in section 2.4.4 via the use of the representation theorem.

We observe that the integrals are defined over the boundaries of the domains, i.e, the interfaces. Provided with these equations and the triangulations of the interfaces (cf. figure 2.6), an approximate solution is obtained. Let us denote by $S = \cup_k S_k$ the union of the surfaces S_k, and E the space of functions square integrable on S. We assume that the surfaces S_k are approximated with a set of n triangles $\{T_i \mid i \in [1, \ldots, n]\}$. Like for the FEM, the solution is approximated in a subspace of finite dimension and this approximation is computed by using test functions.

Just for illustration purposes, let us consider a constant collocation method. The potential is discretized in the subspace of piecewise constant functions, which are constant on each triangle (P0 discretization). The test functions are Dirac functions located at the center of each triangle. We denote the approximation subspace E_h, where h is an index which stands for the size of the largest triangle. A basis $(\phi_i)_i$ of P0 elements for E_h is:

$$\begin{cases} \phi_i(\mathbf{r}) = 1, & \mathbf{r} \in T_i \\ \phi_i(\mathbf{r}) = 0, & \mathbf{r} \notin T_i \end{cases}$$

Any function $f \in E_h$ can then be written $f = \sum_i f_i \phi_i$, where f_i is the constant value of f



(a) The scalp surface extracted from an MRI. (b) An approximation with triangles.

Figure 2.6: Example of triangulated surface used as interface in the boundary element method.

on T_i. Injecting this expression into the equation (2.27) for the electric potential and testing with Dirac functions leads to the problem:

Find $V \in E_h$ such that:

$$\forall i,\ T_i \in S_k,\ \frac{\sigma_k + \sigma_{k+1}}{2} V_i = V_0(\mathbf{r}_i) - \frac{1}{4\pi} \sum_l (\sigma_l - \sigma_{l+1}) \sum_{T_j \in S_l} V_j \int_{T_j} \phi_j(\mathbf{r}') \frac{\mathbf{r}_i - \mathbf{r}'}{\|\mathbf{r}_i - \mathbf{r}'\|^3} \cdot \mathbf{n}_j ds' \ ,$$

where $\mathbf{r}_i$ is the center of triangle T_i and $\mathbf{n}_j$ is the constant normal to triangle T_j. This last equation is of the form

$$V_i = b_i + \sum_j a_{ij} V_j \ ,$$

where the b_i and a_{ij} are constant coefficients that can be computed. The problem takes again the form of a linear system:

$$\mathbf{A}\,[V_i] = \mathbf{b} \ . \tag{2.29}$$

The values V_i of V on each triangle are obtained by resolution of this linear system. This matrix is however singular. The potential is defined up to an additive constant. In order to obtain a non singular matrix another constraint needs to be added. This can be done by forcing the average potential on all surfaces to be zero. This operation is generally referred to as *deflation*. From this electric potential, an approximate solution of the magnetic field generated by the same source can be computed with equation (2.28) [71].

Computationally, one advantage of the BEM is that the matrix $\mathbf{A}$ is generally sufficiently small to use direct methods for the resolution of the linear system. Such methods use factorizations of the matrix $\mathbf{A}$ (*e.g.,* $\mathbf{A} = \mathbf{LU}$) which transform the linear system (2.29) in a new linear system which can be solved very rapidly. Because the contribution of the source $\mathbf{J}^{\mathrm{p}}$ only appears in the right hand side $\mathbf{b}$ of the system (2.29), the factorization of the matrix (which corresponds to the most computationally expensive part) has to be performed once for a given head model, and then the solution of the forward problem can be computed rapidly for many different source distributions.

The procedure described here, uses a piecewise constant discretization for the potential (*i.e.,* P0 elements) and collocation. This is a very coarse level of discretization. In order to improve the numerical precision, it is preferable to discretize the potential with P1 elements,

i.e., in the space of piecewise linear functions, and to use the same P1 functions as test functions [57, 155].

However, the standard BEM derived from Geselowitz's formulas is prone to certain numerical errors. First, if there are large differences between the conductivities of the different compartments of the head model, it can lead to an amplification of the numerical errors [148]. The Isolated Problem Approach can be used to reduce this effect [102]. Second, for sources which are located close to an interface, typically at a distance smaller than the size of the triangles used to describe the surfaces, the accuracy of the BEM drops severely. It is to circumvent this limitation of the BEM that a new formulation has been introduced [131, 132]. This formulation that we are going to present now is called the Symmetric BEM, since its leads to a linear system where the matrix to be inverted is symmetric.

2.4.4 The Symmetric Boundary Element Method (SymBEM)

The Symmetric Boundary Elements Method (SymBEM) is intrinsically a reformulation of the integral equations (2.27) and (2.28) at the origin of the standard BEM. This method is based on advanced representation theorems originally developed in the group of J-C Nédelec [164].

Green Representation Theorem

The Green Representation Theorem states that a piecewise harmonic function can be expressed as a combination of boundary integrals of its discontinuities and the discontinuities of its normal derivative across interfaces.

Let $\partial_{\mathbf{n}} V = \mathbf{n} \cdot \nabla V$ denote the partial derivative of V in the direction of a unit vector $\mathbf{n}$. The restriction of a function f to a surface S_j will be denoted f_{S_j}. We define the discontinuity of a function $f : \mathbb{R}^3 \to \mathbb{R}$ across S_j as

$$[f]_{S_j} = f^-_{S_j} - f^+_{S_j},$$

where the functions f^- and f^+ on S_j are respectively the interior and exterior limits of f:

$$\text{for } \mathbf{r} \in S_j, \qquad f^{\pm}_{S_j}(\mathbf{r}) = \lim_{\alpha \to 0^{\pm}} f(\mathbf{r} + \alpha \mathbf{n}).$$

Let us consider an open region Ω and a function u such that $\Delta u = 0$ in Ω and in $\mathbf{R}^3 \backslash \Omega$. Let $G(\mathbf{r}) = \frac{1}{4\pi \|\mathbf{r}\|}$ be the fundamental solution of the Laplacian such that $-\Delta G = \delta_0$. The Green Representation Theorem states that, for a point $\mathbf{r}$ belonging to $\partial\Omega$,

$$\frac{u^-(\mathbf{r}) + u^+(\mathbf{r})}{2} = -\int_{\partial\Omega} [u]\, \partial_{\mathbf{n}'} G(\mathbf{r} - \mathbf{r}') ds(\mathbf{r}') + \int_{\partial\Omega} [\partial_{\mathbf{n}'} u]\, G(\mathbf{r} - \mathbf{r}') ds(\mathbf{r}') \ .$$

As shown in [131], this representation also holds when Ω is the union of disjoint open sets: $\Omega = \Omega_1 \cup \Omega_2 \cup \ldots \Omega_N$, with $\partial\Omega = S_1 \cup S_2 \cup \ldots S_N$, as in figure 2.7. In this case, for $\mathbf{r} \in S_i$,

$$\frac{u^-(\mathbf{r}) + u^+(\mathbf{r})}{2} = -\sum_{j=1}^{N} \left(\int_{S_j} [u]_{S_j}\, \partial_{\mathbf{n}'} G(\mathbf{r} - \mathbf{r}') ds(\mathbf{r}') + \int_{S_j} [\partial_{\mathbf{n}'} u]_{S_j}\, G(\mathbf{r} - \mathbf{r}') ds(\mathbf{r}') \right) \tag{2.30}$$

The notation is simplified by introducing two integral operators, called the "double-layer" and "single-layer" operators, which map a scalar function f on $\partial\Omega$ to another scalar function on $\partial\Omega$:

$$\big(\mathcal{D} f\big)(\mathbf{r}) = \int\limits_{\partial\Omega} \partial_{\mathbf{n}'} G(\mathbf{r} - \mathbf{r}') f(\mathbf{r}')\, \mathrm{d}s(\mathbf{r}')$$

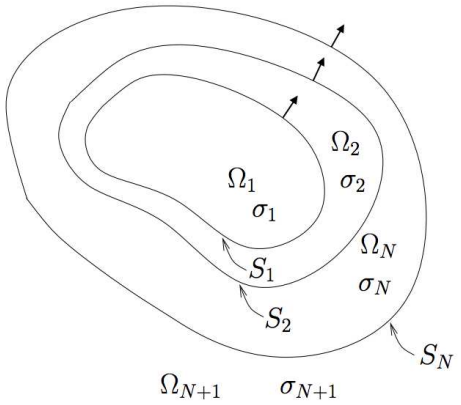


Figure 2.7: The head is modeled as a set of nested regions $\Omega_1, \ldots, \Omega_{N+1}$ with constant isotropic conductivities $\sigma_1, \ldots, \sigma_{N+1}$, separated by interfaces $S_1, \ldots, S_N$. Arrows indicate the normal directions (outward).

and

$$\big(\mathcal{S}f\big)(\mathbf{r}) = \int\limits_{\partial\Omega} G(\mathbf{r}-\mathbf{r}')f(\mathbf{r}')\,\mathrm{d}s(\mathbf{r}') \ .$$

For a given operator $\mathcal{A}$, its restriction which maps a function of S_j to a function of S_i is denoted $\mathcal{A}_{ij}$.

The double-layer BEM

To apply the representation theorem to the forward problem of EEG, a harmonic function which relates the potential and the sources must be produced. Let us decompose the source term as $f = \sum_i f_i$ where the support of each f_i lies inside homogeneous region Ω_i, and consider v_{Ω_i} such that $\Delta v_{\Omega_i} = f_i$ holds in all $\mathbb{R}^3$. The function $v_d = \sum_{i=1}^{N} v_{\Omega_i}$ satisfies $\Delta v_d = f$ and is continuous across each surface S_i, as well as its normal derivative $\partial_{\mathbf{n}} v_d$. The function $u = \sigma V - v_d$ is a harmonic function in Ω, to which (2.30) can be applied. Since $[u]_{S_i} = (\sigma_i - \sigma_{i+1})V_j$ and $[\partial_{\mathbf{n}} u] = 0$, we obtain, on each surface S_i,

$$\frac{\sigma_i + \sigma_{i+1}}{2} V_j + \sum_{j=1}^{N} (\sigma_j - \sigma_{j+1})\,\mathcal{D}_{ij}\,V_j = v_d \ . \tag{2.31}$$

By noticing that:

$$\big(\mathcal{D}f\big)(\mathbf{r}) = \int\limits_{\partial\Omega} f(\mathbf{r}') \frac{\mathbf{r}-\mathbf{r}'}{\|\mathbf{r}-\mathbf{r}'\|^3} \mathrm{d}s(\mathbf{r}')$$

we get that this formula is exactly the formula established by Geselowitz (2.27). Hence, the classical BEM corresponds to a double-layer potential formulation because it involves the double-layer operator $\mathcal{D}$.

An extension of the Green Representation Theorem represents the directional derivative of a harmonic function as a combination of boundary integrals of higher order. This requires two more integral operators: the adjoint $\mathcal{D}^*$ of the double-layer operator, and a hyper-singular operator $\mathcal{N}$ defined by:

$$\big(\mathcal{N}f\big)(\mathbf{r}) = \int\limits_{\partial\Omega} \partial_{\mathbf{n},\mathbf{n}'} G(\mathbf{r}-\mathbf{r}')f(\mathbf{r}')\,\mathrm{d}s(\mathbf{r}') \ .$$

The theorem says that if $\mathbf{r}$ is a point of S_i, then

$$-\frac{\partial_{\mathbf{n}} u^-(\mathbf{r}) + \partial_{\mathbf{n}} u^+(\mathbf{r})}{2} = +\mathcal{N}[u] - \mathcal{D}^*[\partial_{\mathbf{n}} u] \tag{2.32}$$

The Geselowitz formula uses the first boundary integral representation equation (2.30), whereas the Symmetric BEM [131] uses both (2.30) and (2.32) in a formulation combining single- and double-layer potentials.

The symmetric BEM

The originality of the symmetric Boundary Element Method is to consider one piecewise harmonic function per domain: the function u_{Ω_i} equal to $V - \frac{v_{\Omega_i}}{\sigma_i}$ within Ω_i and to $-\frac{v_{\Omega_i}}{\sigma_i}$ outside Ω_i. This function u_{Ω_i} is indeed harmonic in $\mathbb{R}^3 \backslash \partial\Omega_i$, and the representation equations (2.30) and (2.32) can be applied, leading to a system of integral equations involving two types of unknowns: the potential V_i and the normal current $(\sigma \partial_{\mathbf{n}} V)_i$ on each interface.

The surfaces are represented by triangular meshes. To fix ideas, we consider a three-layer geometrical model for the head (cf. figure 2.7). Conductivities of each domain are respectively denoted σ_1, σ_2 and σ_3. The surfaces enclosing these homogeneous conductivity regions are denoted S_1 (inner skull boundary), S_2 (skull-scalp interface) and S_3 (scalp-air interface). Denoting $\psi_i^{(k)}$ the P0 function associated to triangle i on surface S_k, and $\phi_j^{(l)}$ the P1 function associated to node j on surface S_l, the potential V on surface S_k is approximated as $V_{S_k}(\mathbf{r}) = \sum_i x_i^{(k)} \phi_i^{(k)}(\mathbf{r})$, while $p = \sigma \partial_{\mathbf{n}} V$ on surface S_k is approximated by $p_{S_k}(\mathbf{r}) = \sum_i y_i^{(k)} \psi_i^{(k)}(\mathbf{r})$.

As an illustration, considering the source term to be restricted to the brain compartment Ω_1, the variables $\left(\mathbf{x}_k\right)_i = x_i^{(k)}$ and $\left(\mathbf{y}_k\right)_i = y_i^{(k)}$ satisfy the linear system:

$$\begin{bmatrix}
(\sigma_1+\sigma_2)\mathsf{N}_{11} & -2\mathsf{D}_{11}^* & -\sigma_2\mathsf{N}_{12} & \mathsf{D}_{12}^* & 0 \\
-2\mathsf{D}_{11} & (\sigma_1^{-1}+\sigma_2^{-1})\mathsf{S}_{11} & \mathsf{D}_{12} & -\sigma_2^{-1}\mathsf{S}_{12} & 0 \\
-\sigma_2\mathsf{N}_{21} & \mathsf{D}_{21}^* & (\sigma_2+\sigma_3)\mathsf{N}_{22} & -2\mathsf{D}_{22}^* & -\sigma_3\mathsf{N}_{23} \\
\mathsf{D}_{21} & -\sigma_2^{-1}\mathsf{S}_{21} & -2\mathsf{D}_{22} & (\sigma_2^{-1}+\sigma_3^{-1})\mathsf{S}_{22} & \mathsf{D}_{23} \\
0 & 0 & -\sigma_3\mathsf{N}_{32} & \mathsf{D}_{32}^* & \sigma_3\mathsf{N}_{33}
\end{bmatrix}
\begin{bmatrix} \mathbf{x}_1 \\ \mathbf{y}_1 \\ \mathbf{x}_2 \\ \mathbf{y}_2 \\ \mathbf{x}_3 \end{bmatrix} =
\begin{bmatrix} \mathbf{b}_1 \\ \mathbf{c}_1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \tag{2.33}$$

where $\mathbf{b}_1$ (resp. $\mathbf{c}_1$) are the coefficients of the P0 (resp. P1) boundary element decomposition of the source term $\partial_{\mathbf{n}} v_{\Omega_1}$ (resp. $-\sigma_1^{-1} v_{\Omega_1}$).

The blocks N_{ij} and D_{ij} map a potential V_j on S_j to a quantity defined on S_i. The blocks S_{ij} map a normal current p_j on S_j to a quantity defined on S_i. The resulting matrix is block-diagonal, and symmetric, hence the name "symmetric BEM".

With OpenMEEG, the *deflation* is done by forcing the average potential on the external surface to be zero. One can prove that this can be done by correcting the last diagonal block with a matrix filled with ones. With the three layer BEM, this corresponds to the replacement of N_{33} by a matrix $\mathsf{N}_{33} + \alpha \mathbf{1}\mathbf{1}^T$, $\alpha > 0$.

Compared to the standard BEM, the symmetric BEM introduces an additional unknown into the problem: the continuity of the normal current through the interfaces is guaranteed. By doing so the Symmetric BEM leads to larger system matrices but demonstrates significantly higher accuracy than the double-layer BEM [131]. This is illustrated in the next

section where the implementation of the SymBEM we have contributed to develop via the OpenMEEG software project, is compared in terms of precision to other available implementations of the BEM.

2.5 IMPLEMENTATION

We have seen during the presentation of the BEM and Symmetric BEM that the computation of the electric potential produced by a dipole leads to a linear system that we write for simplicity:

$$\mathbf{A}[V] = \mathbf{b} \ .$$

For the SymBEM the notation $[V]$ also contains the discretized normal derivative of the potential, *i.e.,* the normal current.

After applying the deflation to $\mathbf{A}$, the potential (and the normal current with the SymBEM) on each interface is therefore given by:

$$[V] = \mathbf{A}^{-1}\mathbf{b}.$$

In practice, the potential is only measured in EEG at the position of the electrodes on the outer surface of the head model. As a consequence, the forward problem of EEG requires to compute the linear operator, denoted $\mathbf{E}$, that maps the potential defined on the interfaces to the potential at the sensors positions. In practice, this is just an interpolation of the potential computed on the outer surface. This leads to:

$$[V_{eeg}] = \mathbf{E}\mathbf{A}^{-1}\mathbf{b} \ .$$

The forward field, denoted $\mathbf{g}_{eeg}$, for EEG is therefore computed in 5 steps:

- computation of $\mathbf{A}$
- inversion of $\mathbf{A}$
- computation of $\mathbf{b}$
- computation of $\mathbf{E}$
- computation of $\mathbf{g}_{eeg} = \mathbf{E}\mathbf{A}^{-1}\mathbf{b}$

With MEG, one needs to add the computation of the matrix that maps the primary currents directly to the MEG sensors. This corresponds to the $\mathbf{B}_0$ term in equation (2.8). Let us denote by $\mathbf{c}$ this linear operator (considering only one dipolar source). Like with EEG, one also need to compute the operator that computes the effects of the potential on the MEG sensors. The matrix is denoted $\mathbf{D}$.

The forward field, denoted $\mathbf{g}_{meg}$, for MEG is then computed in 6 steps:

- computation of $\mathbf{A}$
- inversion of $\mathbf{A}$
- computation of $\mathbf{b}$
- computation of $\mathbf{D}$
- computation of $\mathbf{c}$
- computation of $\mathbf{g}_{meg} = \mathbf{c} + \mathbf{D}\mathbf{A}^{-1}\mathbf{b}$

When considering multiple dipoles, like in chapter 3 for distributed source models, the forward field for EEG and MEG is computed for each of them and the concatenation of all the

	Programming Language	Licence	DM with spherical HM in M/EEG	MM with spherical HM in MEG	Standard BEM	Standard BEM with ISA enabled	FEM	Symmetric BEM
Brainstorm	Matlab	GPL v2	√	√	√	√		
Simbio	Fortran/C++	GPL v2	√		√	√	√	
Fieldtrip (dipoli)	C	Not open source	√		√	√		
SPM & Fieldtrip (BEMCP)	C/Matlab	GPL v2	√		√			
MNE	C/C++	Not open source	√		√	√		
OpenMEEG	C++	Cecill-B						√

Table 2.1: Review of non commercial software computing the forward problem in M/EEG. DM stands for Dipolar Model, MM for Multipolar Model and HM stands for Head Model.

forward fields leads to what is usually called the leadfield matrix, or the gain matrix, that will be denoted $\mathbf{G}$. Each column of $\mathbf{G}$ is the forward field of one dipole.

2.6 SOFTWARE

2.6.1 Review of non commercial available software

We now present a list of non commercial software packages that can be used to compute the forward problem solutions presented within this chapter. Some of these packages are completely open source: OpenMEEG, Simbio, Fieldtrip (BEMCP implementation) and Brainstorm. The Fieldtrip Toolbox shares its code for M/EEG forward and inverse modeling with the SPM toolbox. It offers two implementations of the BEM. The first one, called *dipoli*, was written by Oostendorp and is not open source (only binary files for Linux are available), while the second one, called *BEMCP*, is opensource and was written by Christoph Phillips during his PhD [176]. The Simbio solver implements a linear Galerkin method (P1 elements with P1 test functions) as described in [57], while *dipoli* and BEMCP implement linear collocation methods. The *dipoli* implementation details can be found in [165]. The MNE Matlab Toolbox also offers a linear collocation implementation of the BEM usable via binary files. This information is summarized in Table 2.1.

We now describe the work accomplished in the OpenMEEG software package. The following paragraph can be read as a short manual for computing the forward problem with OpenMEEG.

2.6.2 OpenMEEG

Let us first present how the description in section 2.5 is transposed in the OpenMEEG naming conventions.

The matrix:

- **A** is called HeadMat or Head Matrix
- $\mathbf{A}^{-1}$ is the inverse of HeadMat
- **B** is called SourceMat or Source Matrix
- **C** is DipSource2MEGMat
- **D** is Head2MEGMat
- **E** is Head2EEGMat

The OpenMEEG package takes as input:

- `subject.geom` : a file describing the geometry of the head (see. table 2.2)
- `subject.cond` : a file containing the conductivity of each tissue of the head (see. table 2.3)
- `eeg_electrodes.txt` : a file containing the 3D positions of each EEG electrode (3 coordinates on each line)
- `dipoles.txt` : a file containing the 3D positions and orientations of each current dipole (6 values on each line)
- `meg_squids.txt` : a file containing the 3D positions and orientations of each MEG sensor (6 values on each line). More complex sensors can also be modeled by integration, or with finite differences like for basic gradiometers.

The different matrices are computed with the following command lines (.bin file extension corresponds to matrix files stored in binary format):

OpenMEEG with the command line

Matrix A:

```
$ om_assemble -HeadMat subject.geom subject.cond HeadMat.bin
```

Note: the abbreviated option names `-HM` or `-hm` can be used instead of `-HeadMat`.

Matrix $\mathbf{A}^{-1}$:

```
$ om_minverser HeadMat.bin HeadMatInv.bin
```

Matrix B:

```
$ om_assemble -DipSourceMat subject.geom subject.cond dipoles.txt
SourceMat.bin
```

Note: the abbreviated option names `-DSM` or `-dsm` can be used instead of `-DSM`.

Matrix E:

```
# Domain Description 1.0

Interfaces 3 Mesh

skull.tri
brain.tri
scalp.tri

Domains 4

Domain Skin 1 -3
Domain Brain -2
Domain Air 3
Domain Skull 2 -1
```

Table 2.2: Sample geometry file for OpenMEEG. It provides the names of the meshes for all the interfaces and the structure by specifying which regions are being separated by each mesh, *e.g.,* the *Skin* region is between the meshes skull.tri (1 means that Skin is outside the first interface on the list, *i.e.,* "skull.tri") and scalp.tri (-3 means that the Skin is inside the third interface on the list, *i.e.,* "scalp.tri").

```
# Properties Description 1.0 (Conductivities)

Air         0.0
Skin        1
Brain       1
Skull       0.0125
```

Table 2.3: Sample conductivity file for OpenMEEG. It specifies the conductivity values for each tissue.

```
$ om_assemble -Head2EEGMat subject.geom subject.cond eeg_electrodes.txt
Head2EEGMat.bin
```

Note: the abbreviated option names `-H2EM` or `-h2em` can be used instead of `-Head2EEGMat`.

Matrix D:

```
$ om_assemble -Head2MEGMat subject.geom subject.cond meg_squids.txt
Head2MEGMat.bin
```

Note: the abbreviated option names `-H2MM` or `-h2mm` can be used instead of `-Head2MEGMat`.

Matrix C:

```
$ om_assemble -DipSource2MEGMat dipoles.txt meg_squids.txt
Source2MEGMat.bin
```

Note: the abbreviated option names `-DS2MM` or `-ds2mm` can be used instead of `-DipSource2MEGMat`.

Matrix $\mathbf{G}_{eeg}$:

```
$ om_gain -EEG HeadMatInv.bin SourceMat.bin Head2EEGMat.bin
GainEEGMat.bin
```

Matrix $\mathbf{G}_{meg}$:

```
$ om_gain -MEG HeadMatInv.bin SourceMat.bin Head2MEGMat.bin
Source2MEGMat.bin GainMEGMat.bin
```

During this PhD we contributed to add a set of features to the OpenMEEG package:

- An scripting interface with Python: A demo script is provided in table 2.4, with its output in table 2.5.
- Parallel processing with OpenMP to speed up computations on machines with multiple processors (it required to rewrite part of the code used by the operators.). Computation times with parallel processing enabled are available in figure 2.10. It can be observed that our parallel implementation offers a significative improvement in terms of computation time. The computation time of $\mathbf{A}$ and $\mathbf{B}$ decreases almost linearly with the number of threads. The matrix inversion is not multithreaded which explains why the computation of $\mathbf{A}^{-1}$ is not improved when increasing the number of threads. With 2 processors, a gain matrix with a realistic head model (cf. figure 2.10(c)) is assembled about 2 times faster, while with 8 processors, the same gain matrix is assembled 3 times faster.
- An advanced testing procedure to guarantee the integrity of the results obtained by the forward problem computation. Results can be compared at no cost to analytical solutions obtained with 3 layer spherical head models like the one presented in figure 2.8. This procedure is based on the *CTest* testing software. The output of the testing procedure is presented in table 2.6.
- A Matlab interface within the Fieldtrip Toolbox and SPM Toolbox (see table 2.13).
- A multi platform packaging system based on *CPack* allowing easy deployment on all architectures (Linux, Mac and Windows environments).

Thanks to the integration of the OpenMEEG software into the Fieldtrip Toolbox, we have been able to demonstrate that the precision obtained by our numerical solution of the forward problem clearly outperforms the standard BEM implementations offered by the SPM and Fieldtrip Toolbox (dipoli and BEMCP). We have also been implicated in the development

```
#!/usr/bin/env python

import openmeeg as om

# =============
# = Load data =
# =============

condFile='om_demo.cond'
geomFile='om_demo.geom'
dipoleFile='cortex.dip'
squidsFile='meg_squids.txt'
electrodesFile='eeg_electrodes.txt'

geom = om.Geometry()
geom.read(geomFile,condFile)

dipoles = om.Matrix()
dipoles.load(dipoleFile)

squids = om.Sensors()
squids.load(squidsFile)

electrodes = om.Matrix()
electrodes.load(electrodesFile)

# ================================================
# = Compute forward problem (Build Gain Matrices) =
# ================================================

gaussOrder = 3; # Integration order over the triangles in the BEM

hm       = om.HeadMat(geom,gaussOrder)
hminv    = hm.inverse()
dsm      = om.DipSourceMat(geom,dipoles,gaussOrder)
ds2mm    = om.DipSource2MEGMat(dipoles,squids)
h2mm     = om.Head2MEGMat(geom,squids)
h2em     = om.Head2EEGMat(geom,electrodes)
gain_meg = om.GainMEG(hminv,dsm,h2mm,ds2mm)
gain_eeg = om.GainEEG(hminv,dsm,h2em)

print "hm          : %d x %d"%(hm.nlin(),hm.ncol())
print "hminv       : %d x %d"%(hminv.nlin(),hminv.ncol())
print "dsm         : %d x %d"%(dsm.nlin(),dsm.ncol())
print "ds2mm       : %d x %d"%(ds2mm.nlin(),ds2mm.ncol())
print "h2mm        : %d x %d"%(h2mm.nlin(),h2mm.ncol())
print "h2em        : %d x %d"%(h2mm.nlin(),h2mm.ncol())
print "gain_meg    : %d x %d"%(gain_meg.nlin(),gain_meg.ncol())
print "gain_eeg    : %d x %d"%(gain_eeg.nlin(),gain_eeg.ncol())
```

Table 2.4: Demo script for computing the forward problem with OpenMEEG in Python.

```
Sorted List :  1 0 2
Sorted Domains :  Brain Skull Scalp Air
Total number of points    : 126
Total number of triangles : 240

Checking
Mesh 0 : internal conductivity = 1 and external conductivity = 0.0125
Mesh 1 : internal conductivity = 0.0125 and external conductivity = 1
Mesh 2 : internal conductivity = 1 and external conductivity = 0
OPERATOR S...
[********************]
OPERATOR S...
[********************]
OPERATOR S...
[********************]
OPERATOR N...
[********************]
OPERATOR N...
[********************]
OPERATOR N...
[********************]
OPERATOR D (Optimized)...
[********************]
OPERATOR D (Optimized)...
[********************]
OPERATOR D (Optimized)...
[********************]
OPERATOR D (Optimized)...
[********************]
OPERATOR S...
[********************]
OPERATOR S...
[********************]
OPERATOR N...
[********************]
OPERATOR N...
[********************]
OPERATOR D (Optimized)...
[********************]
[********************]
[********************]
hm          : 286 x 286
hminv       : 286 x 286
dsm         : 286 x 42
ds2mm       : 162 x 42
h2mm        : 162 x 286
h2em        : 162 x 286
gain_meg    : 162 x 42
gain_eeg    : 42 x 42
```

Table 2.5: Output of Python demo script presented in table 2.4.

```
Running tests...
Start processing tests
Test project openmeeg_trunk
  1/ 73 Testing matlibtest                          Passed
  2/ 73 Testing HM-Head1                            Passed
  3/ 73 Testing HMINV-Head1                         Passed
  4/ 73 Testing SSM-Head1                           Passed
  5/ 73 Testing AI-Head1                            Passed
  6/ 73 Testing H2EM-Head1                          Passed
  7/ 73 Testing SurfGainEEG-Head1                   Passed
  8/ 73 Testing ESTEEG-Head1                        Passed
  9/ 73 Testing EEG-HEAT-Head1                      Passed
 10/ 73 Testing EEG-MN-Head1                        Passed
 11/ 73 Testing EEG-TV-Head1                        Passed
 12/ 73 Testing H2MM-Head1                          Passed
 13/ 73 Testing SS2MM-Head1                         Passed
 14/ 73 Testing SurfGainMEG-Head1                   Passed
 15/ 73 Testing ESTMEG-Head1                        Passed
 16/ 73 Testing MEG-HEAT-Head1                      Passed
 17/ 73 Testing MEG-MN-Head1                        Passed
 18/ 73 Testing MEG-TV-Head1                        Passed
 19/ 73 Testing DSM-Head1                           Passed
 20/ 73 Testing DS2MM-Head1                         Passed
 21/ 73 Testing DipGainEEG-Head1                    Passed
 22/ 73 Testing DipGainMEG-Head1                    Passed

 ...

 52/ 73 Testing compareEEGEST-dip-Head1-d1          Passed
 53/ 73 Testing compareEEGEST-dip-Head2-d1          Passed
 54/ 73 Testing compareEEGEST-dip-Head1-d2          Passed
 55/ 73 Testing compareEEGEST-dip-Head2-d2          Passed
 56/ 73 Testing compareEEGEST-dip-Head1-d3          Passed
 57/ 73 Testing compareEEGEST-dip-Head2-d3          Passed
 58/ 73 Testing compareEEGEST-dip-Head1-d4      ***Failed  - supposed to fail
 59/ 73 Testing compareEEGEST-dip-Head2-d4          Passed
 60/ 73 Testing compareEEGEST-dip-Head1-d5      ***Failed  - supposed to fail
 61/ 73 Testing compareEEGEST-dip-Head2-d5          Passed
 62/ 73 Testing compareMEGEST-dip-Head1-d1          Passed
 63/ 73 Testing compareMEGEST-dip-Head2-d1          Passed
 64/ 73 Testing compareMEGEST-dip-Head1-d2          Passed
 65/ 73 Testing compareMEGEST-dip-Head2-d2          Passed
 66/ 73 Testing compareMEGEST-dip-Head1-d3          Passed
 67/ 73 Testing compareMEGEST-dip-Head2-d3          Passed
 68/ 73 Testing compareMEGEST-dip-Head1-d4          Passed
 69/ 73 Testing compareMEGEST-dip-Head2-d4          Passed
 70/ 73 Testing compareMEGEST-dip-Head1-d5          Passed
 71/ 73 Testing compareMEGEST-dip-Head2-d5          Passed
 72/ 73 Testing compareMEGEST-dip-Head1-d6      ***Failed  - supposed to fail
 73/ 73 Testing compareMEGEST-dip-Head2-d6          Passed

100% tests passed, 0 tests failed out of 73
```

Table 2.6: Output of testing procedure for OpenMEEG. Output is systematically compared to analytical solutions with spherical head models.

of the SimBio package during this PhD allowing to add to the comparison the SimBio implementation of the BEM with ISA. The sample dataset used to demonstrate this is presented in figure 2.8 with 5 dipoles at various distances from the inner layer. The quantification of performance is based on the Relative Difference Measure (RDM) and the ratio of Magnitude (MAG) between each numerical solution and the analytical solution. The analytical solutions are computed with the formulas detailed in section 2.3.1.

The RDM between two forward fields is defined as:

$$RDM(g_{\text{numeric}}, g_{\text{analytic}}) = \left\| \frac{g_{\text{numeric}}}{\|g_{\text{numeric}}\|} - \frac{g_{\text{analytic}}}{\|g_{\text{analytic}}\|} \right\| \in [0\ ,\ 2] \ .$$

The closer to 0 is the RDM, the better it is.

The MAG between two forward fields is defined as:

$$MAG(g_{\text{numeric}}, g_{\text{analytic}}) = \frac{\|g_{\text{numeric}}\|}{\|g_{\text{analytic}}\|}$$

The closer to 1 is the MAG, the better it is.

The results with EEG forward fields are presented in figure 2.9 for three-shell spherical head models having 3 different point samplings on each interface. One with only 42 vertices per interface and 42 EEG electrodes, one with 162 points per interface and 162 EEG electrodes, and one with 642 points per interface and 642 EEG electrodes. The radii of the 3 shells, supposed to reproduce the inner surface of the skull, the outer surface of the skull and the skin, were set to 88, 92 and 100. The conductivities of the three domains were set to the ones commonly used in the literature: 1, 1/80 and 1.

From these simulations we can conclude that:

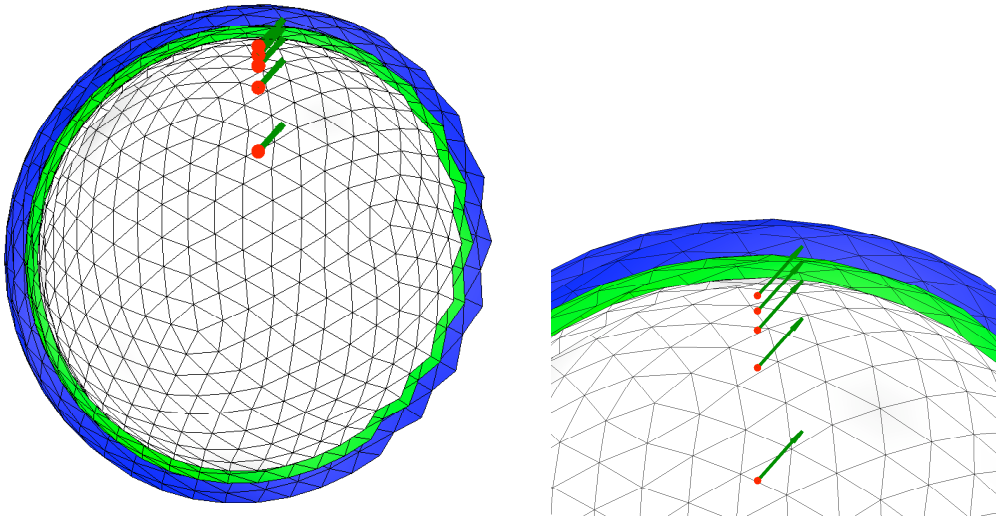
- The BEMCP implementation is clearly the less accurate solver.
- SimBio and DIPOLI give very similar results.
- Our implementation of the Symmetric BEM provides the most accurate solutions.

The numerical values plotted in figure 2.9 are reproduced in tables 2.7, 2.8, 2.9, 2.10, 2.11 and 2.12.

Remark. Late in our investigations, we also experimented the Matlab BEM Toolbox developed by Matti Stenroos (avalaible at `http://peili.hut.fi/BEM/`). This software implements the classical formulation of the BEM from Geselowitz with a linear collocation method. Our first results with this toolbox demonstrate that it does not provide a better accuracy than OpenMEEG either.

Distance to inner layer	45.5	20	11.5	7.25	3.85
BEMCP	1.87e-01	1.06e+00	1.77e+00	1.84e+00	1.86e+00
DIPOLI	1.12e-01	2.58e-01	5.12e-01	6.51e-01	7.36e-01
OpenMEEG	4.23e-02	9.99e-02	1.57e-01	2.03e-01	2.45e-01
SimBio	5.89e-02	1.88e-01	4.44e-01	5.63e-01	6.23e-01

Table 2.7: RDMs precision results with 42 vertices per interface.



(a) 3 layers spherical head model (b) Zoom

Figure 2.8: Spherical head model with 5 dipoles close to the inner layer.

Distance to inner layer	45.5	20	11.5	7.25	3.85
BEMCP	1.59e-01	3.69e-01	1.44e+00	1.82e+00	1.86e+00
DIPOLI	5.45e-02	9.77e-02	1.34e-01	2.89e-01	5.59e-01
OpenMEEG	2.55e-02	5.23e-02	8.10e-02	1.03e-01	1.29e-01
SimBio	4.49e-02	7.73e-02	1.14e-01	2.77e-01	5.46e-01

Table 2.8: RDMs precision results with 162 vertices per interface.

Distance to inner layer	45.5	20	11.5	7.25	3.85
BEMCP	6.38e-02	1.83e-01	2.77e-01	6.62e-01	1.83e+00
DIPOLI	2.39e-02	4.32e-02	5.25e-02	5.97e-02	1.76e-01
OpenMEEG	2.76e-03	6.91e-03	1.03e-02	1.31e-02	1.77e-02
SimBio	1.91e-02	3.43e-02	4.15e-02	4.88e-02	1.79e-01

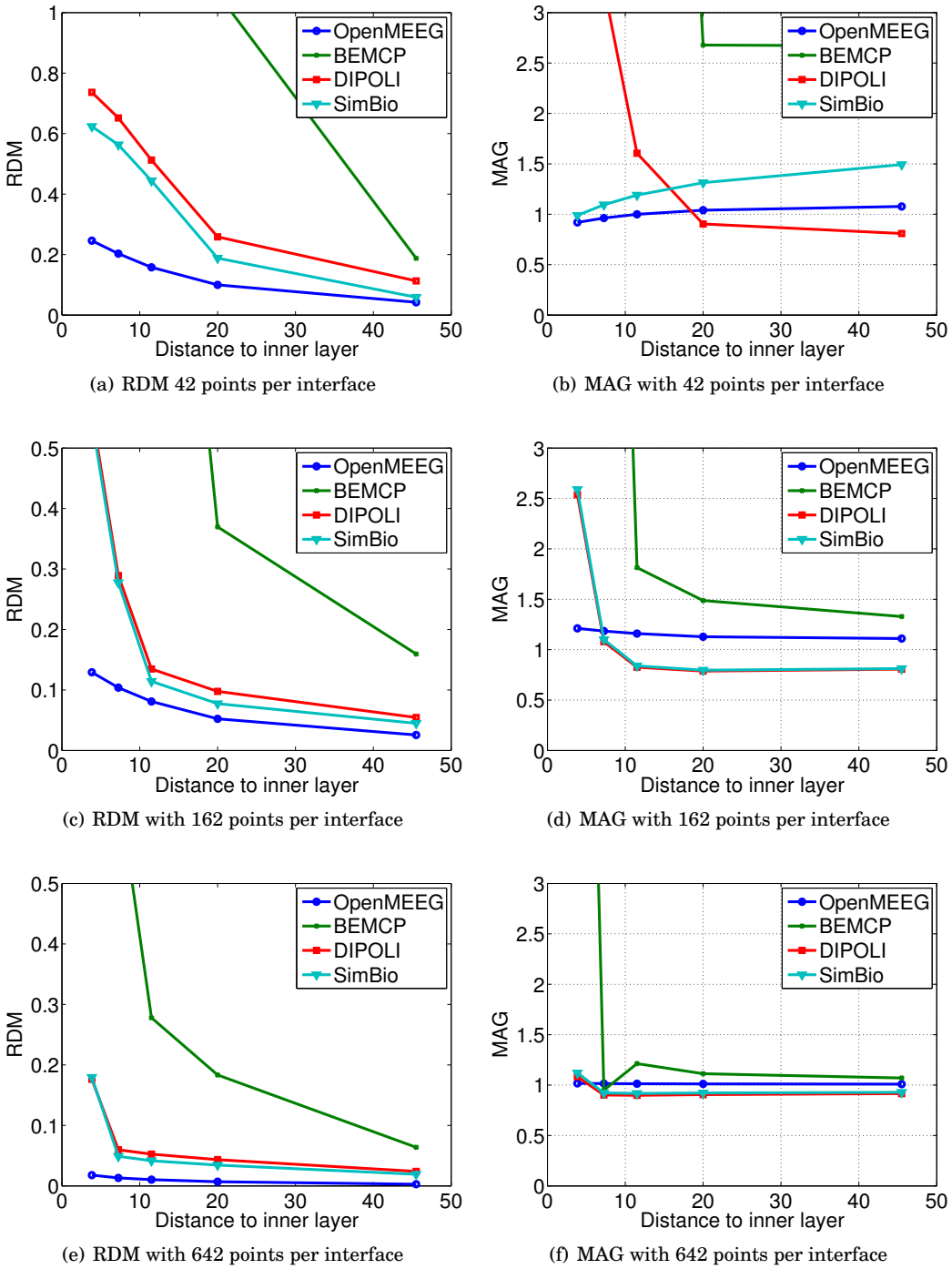
Table 2.9: RDMs precision results with 642 vertices per interface.

Distance to inner layer	45.5	20	11.5	7.25	3.85
BEMCP	2.66e+00	2.67e+00	1.44e+01	5.27e+01	2.40e+02
DIPOLI	8.09e-01	9.03e-01	1.60e+00	3.25e+00	9.45e+00
OpenMEEG	1.07e+00	1.04e+00	9.99e-01	9.62e-01	9.19e-01
SimBio	1.49e+00	1.31e+00	1.18e+00	1.09e+00	9.88e-01

Table 2.10: MAGs precision results with 42 vertices per interface.

Distance to inner layer	45.5	20	11.5	7.25	3.85
BEMCP	1.32e+00	1.48e+00	1.81e+00	1.16e+01	6.94e+01
DIPOLI	8.06e-01	7.87e-01	8.25e-01	1.07e+00	2.53e+00
OpenMEEG	1.11e+00	1.12e+00	1.15e+00	1.18e+00	1.21e+00
SimBio	8.11e-01	7.96e-01	8.38e-01	1.09e+00	2.58e+00

Table 2.11: MAGs precision results with 162 vertices per interface.



(a) RDM 42 points per interface

(b) MAG with 42 points per interface

(c) RDM with 162 points per interface

(d) MAG with 162 points per interface

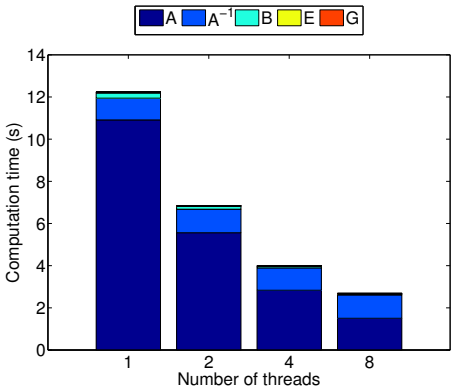
(e) RDM with 642 points per interface

(f) MAG with 642 points per interface

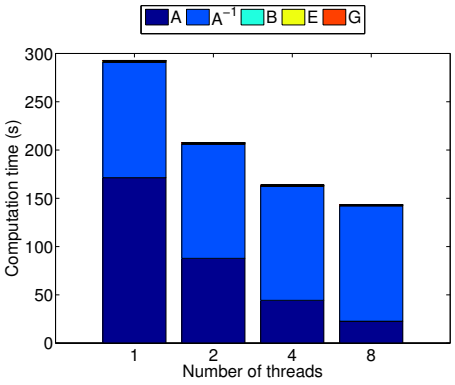
Figure 2.9: Evaluation of precision of different implementations of the BEM with three layers spherical head models. We observe that the Symmetric BEM outperforms in term of precision the other methods.

Distance to inner layer	45.5	20	11.5	7.25	3.85
BEMCP	1.07e+00	1.11e+00	1.21e+00	9.49e-01	1.11e+01
DIPOLI	9.15e-01	9.05e-01	8.98e-01	9.01e-01	1.08e+00
OpenMEEG	1.00e+00	1.01e+00	1.01e+00	1.01e+00	1.01e+00
SimBio	9.28e-01	9.21e-01	9.16e-01	9.22e-01	1.12e+00

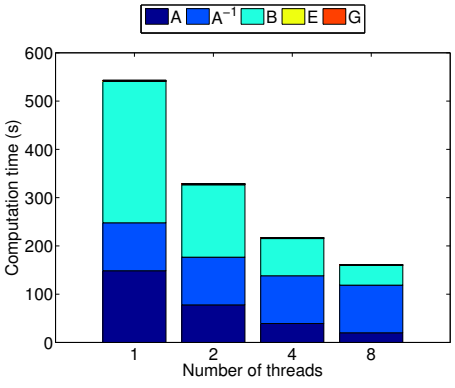
Table 2.12: MAGs precision results with 642 vertices per interface.



(a) Head model with 3 spheres (162 points per interface) and 1 dipole.



(b) Head model with 3 spheres (642 points per interface) and 1 dipole.



(c) Realistic head model with 600, 638 and 625 points on the 3 interfaces and 14055 dipoles.

Figure 2.10: Computation times with parallel processing enabled for all the steps required to compute an EEG leadfield. Tests on 3 head models: 2 spherical head models with 1 dipolar source and 3 layers (162 and 642 vertices per interface) and 1 realistic head model with 600, 638 and 625 points on the 3 interfaces and 14055 dipolar sources. Computation was performed on a Linux 64bit architecture with 8 processors and 64 GB of RAM. Thanks to the parallel implementation of the operators, the computation time of $\mathbf{A}$ and $\mathbf{B}$ decreases almost linearly with the number of threads. The matrix inversion is not multithreaded which explains why the computation of $\mathbf{A}^{-1}$ is not improved when increasing the number of threads.

```
% ==========================================
% = Generate a 3 layers spherical headmodel =
% ==========================================

% 3 Layers
r = [100 92 88]; % radius of each interface
c = [1 1/80 1]; % conductivity within each interface

[pnt,tri] = icosahedron162; % sphere with 162 vertices per interface

% create a set of electrodes on the outer surface
sens.pnt = max(r) * pnt;
sens.label = {};
nsens = size(sens.pnt,1);
for ii=1:nsens
  sens.label{ii} = sprintf('vertex%03d', ii);
end

% Position of the dipole
pos = [0 0 70];

% create a BEM volume conduction model (3 nested interfaces)
vol = [];
for ii=1:length(r)
    vol.bnd(ii).pnt = pnt * r(ii);
    vol.bnd(ii).tri = tri;
end
vol.cond = c;

% ========================
% = Compute the leadfield =
% ========================

% compute the BEM
cfg.method = 'openmeeg'; % can be dipoli or bemcp
vol = prepare_bemmodel(cfg, vol);
lf_openmeeg = compute_leadfield(pos, sens, vol);

% lf_openmeeg is a 162 x 3 matrix
% each column of lf_openmeeg is the forward field of a dipole
% in one direction of the coordinate system
```

Table 2.13: Computing an EEG leadfield with Fieldtrip and OpenMEEG.

2.7 CONCLUSION

Brain functional imaging with M/EEG requires an efficient and accurate forward model. In this section, we have presented the general framework to achieve good forward modeling. It implied to introduce the physics with Maxwell equations, to make a set of hypothesis like the quasi-static approximation and to use simplified head models. We have presented the theory and discussed some implementation details before finally demonstrating that the forward modeling that we contributed to develop and promote in the M/EEG community is the most accurate BEM solver available.

Before closing this chapter, we would like to mention recent and promising work on forward modeling with anisotropic conductivity models [170], *i.e.,* models where the conductivity can vary inside the same tissue. This method avoids the meshing step required by the above FEM and BEM methods and provides quite accurate solutions in a very reasonable amount of time.

We will now move on to the other fundamental aspect of brain functional imaging with MEG: the inverse problem.

CHAPTER 3

THE INVERSE PROBLEM WITH DISTRIBUTED SOURCE MODELS

In chapter 2, we have seen how the tiny electromagnetic fields produced by the neural activity is modeled to understand what is measured with M/EEG devices. This aspect of the processing of M/EEG data is called the *forward problem*. Its counterpart is the *inverse problem*. The inverse problem of M/EEG is the procedure that consists in recovering the distribution of the neural generators that have produced the measurements. Three main types of approaches exist to solve this problem:

1. The parametric models usually referred to as *dipole fitting* approaches.
2. The *beamforming* or *scanning* techniques.
3. The *image-based methods* with distributed source models.

In this chapter, the first two approaches are briefly explained and commented. The last one, the *image-based* method, is presented in more detail. The M/EEG inverse problem is presented in a classical framework where the solution is penalized with a regularization prior. Standard priors are based on ℓ_2 norms, leading to differentiable optimization problems and closed-form solutions. In this chapter, focus is put on such priors, going from the simple "Minimum-Norm" (MN) approach to spatiotemporal solvers and learning-based methods recently presented in the literature.

This chapter covers the methodological aspects of multiple inverse solvers based on ℓ_2 priors. It presents for each of them the hypotheses made and the eventual limitations when used with M/EEG data. The optimization strategies employed will also be commented and discussed. For experimental results and simulation studies with the reviewed solvers, we refer the reader to the original papers. We also provide the code snippets to use these solvers using EMBAL , an open source toolbox that we wrote during this thesis.

Contents

3.1 General introduction to inverse methods

Notations

All matrices and vectors are written in bold letters. Matrices, such as $\mathbf{A}$, are written in upper case, whereas vectors, such as $\mathbf{b}$, are in lower case. A real valued matrix $\mathbf{A} \in \mathbb{R}^{n \times p}$ has n rows and p columns. The notation $\|\mathbf{A}\|_F$ stands for the Frobenius norm of $\mathbf{A}$, while $|||\mathbf{A}|||_2$, or simply $|||\mathbf{A}|||$, stands for the spectral norm. The notation $\|\mathbf{b}\|_2$ stands for the ℓ_2 norm of the vector $\mathbf{b}$. The matrix $\mathbf{I}$ stands for the identity matrix.

3.1.1 Parametric models and dipole fitting approaches

The dipole fitting approaches assume that the measured data have been produced by a small number of active regions that can each be modeled by an equivalent current dipole (ECD). The number of ECDs, denoted K, is fixed. Each dipole i has a position $\mathbf{r}_i$ and a moment $\mathbf{q}_i$. The strength of the dipole is given by $x_i = \|\mathbf{q}_i\|_2$. We denote $\mathbf{g}_i(\mathbf{r}_i, \mathbf{q}_i)$ the forward field produced by this dipole and $\mathbf{m}$ the M/EEG measurements (cf. chapter 2). Since the forward field depends linearly on x_i, the forward field $\mathbf{g}_i(\mathbf{r}_i, \mathbf{q}_i)$ can be rewritten $\mathbf{g}_i(\mathbf{r}_i, \frac{\mathbf{q}_i}{\|\mathbf{q}_i\|})x_i$. The data $\mathbf{m}$ can correspond either to one time instant or to a block of time samples. If dipoles are allowed to move during the time window of interest, the method is called *moving dipole* whereas if it can only rotate it is referred to as *rotating dipole*. Parametric dipole fitting algorithms minimize a data fit cost function such as the Frobenius norm of the residual [11, 156, 193]:

$$\min_{(\mathbf{r}_i, \mathbf{q}_i)_{i=1,\dots,K}} \left\| \mathbf{m} - \sum_{i=1}^{K} \mathbf{g}_i \left(\mathbf{r}_i, \frac{\mathbf{q}_i}{\|\mathbf{q}_i\|} \right) x_i \right\|_F^2 .$$

This optimization problem is non linear, and solvers are easily trapped in local minima as soon as $K > 1$. The optimization strategies employed range from Levenberg-Marquardt and Nelder-Meade downhill simplex searches to global optimization schemes using multistart methods, genetic algorithms and simulated annealing [209].

The main limitation with these methods is that the user has to fix a priori the number of active regions. For these reasons, dipole fitting approaches are commonly used with only one dipole, or sometimes a few, after one is set to a known position. These limitations imply that such methods can only be used reliably with one very focal active region. This is usually a valid assumption for brain activations occurring shortly after stimulation. However, when other functional imaging data with high spatial resolution, like fMRI activation maps, are available, the position of the dipoles can be supposed as known. In this case, multiple dipoles can be manually positioned. Their amplitudes and eventually their orientations are the only estimated parameters. Such an approach is illustrated in chapter 5 where investigations on the human visual cortex with this methodology are presented.

3.1.2 Scanning methods: the beamformers

The Bearmforming approaches, a.k.a., scanning methods, avoid the convexity issue by scanning a region of interest, typically the gray matter forming the cortical mantle. It is also possible to sample dipoles on a regular grid within the cortical envelope. An estimator of the contribution of each putative source location to the data can be derived either via spatial filtering techniques or signal classification indices. Historically scanning methods were first introduced in the radar and sonar community.

```
clear options
% C is the covariance matrix
options.C = C;
% pct a percentage to regularize the inversion of C
options.pct = 10;
[X,W] = lcmv_inverse(M,G,options);
% W contains the spatial filters of all the dipoles
```

Table 3.1: Running a LCMV beamformer with EMBAL .

In its simplest presentation, a spatial filter is a vector $\mathbf{w}$, function of the location and orientation of the dipole of interest, that when correlated to the measurements $\mathbf{m}$ provides an estimate of the moment's amplitude. Equivalently:

$$x(\mathbf{r}) = \mathbf{w}^T\mathbf{m}$$

where x denotes the moment's amplitude of the dipolar source considered. A well designed spatial filter, should *filter out* sources which do not come from a small volume around $\mathbf{r}$.

When considering dipolar sources with unconstrained orientation, the spatial filter $\mathbf{w}$ is a matrix with 3 columns (one per orientation) and the result of $\mathbf{w}^T\mathbf{m}$ provides the moment of the dipole.

The simplest spatial filter is called a *matched filter*. Let $\mathbf{g}_i$ be the forward field of a dipole at position $\mathbf{r}_i$ with a normalized moment $\mathbf{q}$ ($\|\mathbf{q}\|_2 = 1$). The vector $\mathbf{g}_i$ can be seen as the ith column of the lead field and the *matched filter* is obtained by normalizing this column. The spatial filter for this position and orientation is given by

$$\mathbf{w} = \frac{\mathbf{g}_i}{\|\mathbf{g}_i\|_2}$$

This approach guarantees that, when only one source is active, the absolute maximum of the estimate corresponds to the true maximum. In practice this assumption is usually not valid. Since the correlation between the columns of the leadfield matrix is high, the spatial resolution of the matched filter is limited. The goal of more advanced spatial filters is to estimate the activity at a source point while avoiding the crosstalk from other regions. By doing so, the perturbing sources have little influence on the estimation in the region of interest.

The most common spatial filter is the Linearly Constrained Minimum Variance (LCMV) beamformer [215]. It attempts to minimize the beamformer output power subject to a unity gain constraint:

$$\mathbf{w}_{LCMV} = \arg\min_{\mathbf{w}} \operatorname{trace}(\mathbf{w}^T\mathbf{C}\mathbf{w}) \text{ subject to } \mathbf{w}^T\mathbf{g} = 1 \tag{3.1}$$

where $\mathbf{C}$ is the data covariance matrix. By constraining the gain at the considered location and minimizing the energy projected from elsewhere the LCMV beamformer limits the influence of the noise and the crosstalk between the different sources. The constrained optimization problem (3.1) is solved with the method of Lagrange Multipliers under the assumptions of decorrelation between different sources, and between the sources and the noise. The solution is given by:

$$\mathbf{w}_{LCMV}^T = (\mathbf{g}^T\mathbf{C}^{-1}\mathbf{g})^{-1}\mathbf{g}^T\mathbf{C}^{-1} \ .$$

The formula shows it is important to have a correct estimate of the covariance matrix $\mathbf{C}$ which implies in practice to have a sufficient amount of data.

Synthetic aperture magnetometry (SAM) [217] is an alternative the LCMV beamformer. Contrary to LCMV, SAM works on statistical quantities based on the differences between

```
clear options
% k is the dimension of the signal subspace
options.k = k;
[X] = music_inverse(M,G,options);
% X contains the scores obtained by the MUSIC cost function
```

Table 3.2: Running MUSIC with EMBAL .

a control period and the period of activations. Therefore SAM integrates some additional information linked to the design of the experimental paradigms.

An attractive feature of the beamformer methods is that they do not require any a priori on the number of underlying sources. However, they make the strong assumption that the activations of the different sources are uncorrelated. This hypothesis is a particularly critical point. Neural activations in different parts of the brain often co-activate, forming what a network of correlated sources. Even if simulation results [194] and evaluation on real data [97] seem to indicate LCMV-based beamforming methods are robust to moderate levels of source/interference correlation, it is still a fundamental limitation of spatial filtering methods.

The alternative to spatial filtering originates also from the radar and sonar community. These methods are based on signal classification between signal and noise via signal subspaces. The MUltiple SIgnal Classification (MUSIC) is the most popular of these methods [154]. In the MUSIC algorithm the space spanned by the measurements $\mathbf{m}$ is divided with an SVD between the signal space and the noise space. If we write the SVD of $\mathbf{m} = \mathbf{USV}^T$ and estimate the rank of the signal space to r, the signal space is spanned by the r first columns of $\mathbf{U}$ denoted $\mathbf{U}_r$. Note that for the SVD to make sense, the measurements $\mathbf{m}$ have to contain multiple time instants. The cost function associated to MUSIC in the beamforming framework is given by:

$$\frac{\|(\mathbf{I} - \mathbf{U}_r\mathbf{U}_r^T)\mathbf{g}_i\|}{\|\mathbf{g}_i\|} \quad . \tag{3.2}$$

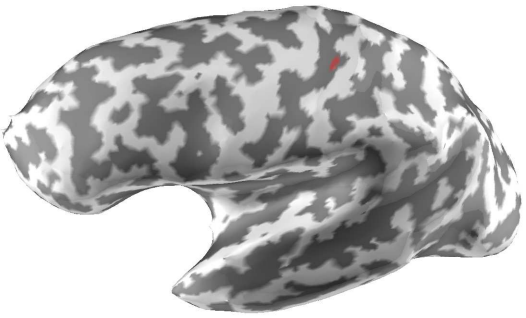
The linear operator $(\mathbf{I} - \mathbf{U}_r\mathbf{U}_r^T)$ acts as an orthogonal projection onto the noise subspace. The smaller is this score the more the ith dipole contributes to the measurements.

A greedy strategy that aims at modeling source configurations with multiple generators has been inspired by the MUSIC algorithm. This variant is called Recursively APplied MUSIC (RAP MUSIC) [153] since it consists in applying the MUSIC cost function successively after removing the contribution of the previously identified sources. Like matching pursuit algorithms are used for sparse signal decomposition over dictionary of atoms [144], the RAP MUSIC method adopts a greedy strategy to select the relevant dipoles in a dictionary of sources.

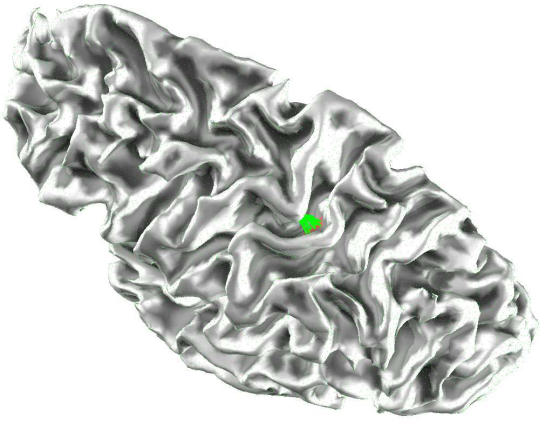
The MUSIC algorithm is relatively popular in the M/EEG community, probably because of its robustness to noise and its ability to present precise locations for the current generators. However, the necessity to set a priori the size of the signal subspace can be an issue. When the amplitudes of the singular values of $\mathbf{m}$ do not exhibit a sharp drop-down after singular value r demonstrating that $\mathbf{m}$ can be approximated with a rank r matrix, the definition of the rank is left to the experience of the user. One clear advantage of the MUSIC method over the LCMV beamformer is the relaxation of the decorrelation constraint between the sources. However, MUSIC requires that the active dipoles have linearly independent time series.

3.1.3 Image-based methods

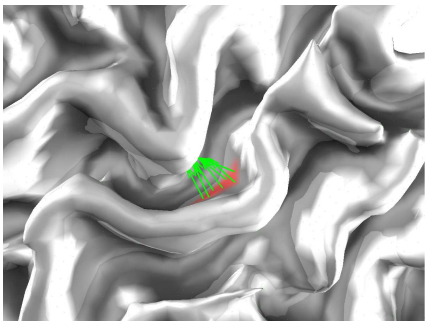
The alternative to dipole fitting and scanning methods is the *image-based* approach. With this method source models, typically dipoles, are sampled, or distributed, over the source space. The source space can be defined as a volume or as a surface. When considering dipolar source models and a volumetric source space, dipoles are typically sampled on a regular 3D grid within the brain region. With surface-based source models, the dipoles are typically distributed over the cortical mantle modeled as a triangular mesh. Such an approach was pioneered by Dale and Sereno in [50], who presented surface-based reconstructions of the neural activations. The neural activations are represented as scalar valued data, hence the comparison with images. The difference comes from the discrete space over which the activations are defined. Rather than having to deal with scalar valued data defined on a 2D or 3D grid like in standard image processing applications, the space is a surface tessellation defined with vertices and triangles. The surface-based distributed source model is illustrated on a synthetic dataset presented in figure 3.1.



(a) Active region on inflated cortical mesh



(b) Active dipoles with constrained orientations



(c) Zoom on active dipoles with constrained orientations

Figure 3.1: Surface-based distributed dipolar sources illustration. The synthetic active region is located on the posterior part of the central sulcus where stands the human primary somatosensory cortex (S1).

Source reconstructions on the cortical surface require to segment the cortex using a T1

MRI image (this type of anatomic data was used to estimate the triangulated interfaces for the BEM in section 2.4.3). This step is rather complex but well handled by software such as BrainVisa [38] or FreeSurfer [51], which provide almost fully automatic pipelines to run the segmentations. Such pipelines are generally not integrated in commercial M/EEG source imaging software[1] that therefore only provide volumetric source spaces with 3D grids.

The current generators that produce the electromagnetic fields are known to be located in the gray matter forming the cortex. This implies that the estimated sources should at least be constrained to be located within the gray matter. This is achieved with surface source models. To argue even more in favor for such models, we would like to mention that the fMRI community also tends to map the 3D data acquired onto cortical surfaces [66]. Another reason for this is that anatomical landmarks are more easily defined on cortical segmentation than on volumetric data.

Orientation vs. no orientation constraint

With distributed dipolar source models, the orientation of the dipoles can either be defined a priori using the normal to the cortical mesh (cf. figure 3.1(c)), or left unconstrained. When the dipoles orientations are left unconstrained, 3 orthogonal dipoles are positioned at each location. With MEG, since sensors are blind to the radial component of the field, only 2 can be used. Considering our knowledge on the structure of the neural assemblies formed by the pyramidal neurons (cf. chapter 1), constraining the orientation is a reasonable assumption. One can also argue that the more a priori are used to compute neural estimates, the better it is. However, practice shows that the orientation is a critical parameter for a dipole since it affects its forward field on the M/EEG sensors a lot more that its 3D position. This suggests that if orientation constraints are used, the normals to the cortical mesh should be very accurately estimated. Depending on the brain location of the sources this can be more or less challenging.

In this chapter, many illustrations are presented on the somatosensory cortex lying on the post-central gyrus. The central sulcus and central gyrus of the cortex are major structures of the human cortex and are very well segmented with anatomical pipelines. For this reason the orientation constraint is generally well justified in this brain region.

3.2 MINIMUM NORM SOLUTIONS AND ITS VARIANTS

When orientations are fixed and only the amplitudes of the dipolar current generators need to be estimated, the forward problem results in the following linear problem:

$$\mathbf{M} = \mathbf{G}\mathbf{X} + \mathbf{E} \tag{3.3}$$

where G stands the forward operator, M corresponds to the measurements (Electric potential or/and magnetic field), $\mathbf{X}$ contains the unknown amplitudes of the sources and $\mathbf{E}$ is the noise.

We denote the number of sources by d_x, the number of sensors by d_m and the number of time instants by d_t. With these notations, we have, $\mathbf{M} \in \mathbb{R}^{d_m \times d_t}$, $\mathbf{G} \in \mathbb{R}^{d_m \times d_x}$, $\mathbf{X} \in \mathbb{R}^{d_x \times d_t}$ and $\mathbf{E} \in \mathbb{R}^{d_m \times d_t}$.

In practice, d_m is in the range of 10, for low resolution EEG, and 400, for high resolution MEG and EEG combined studies. The parameter d_t is commonly between 1 and a few thousand. With the digital amplifiers used in M/EEG, the sampling rate can be over 1000 Hz

[1] To our knowledge only the software package Curry from `http://www.neuroscan.com/` provides segmentations and surface-based reconstructions.

which leads to high values of d_t when recording several seconds of signal. The number of sources d_x is given by the number of dipoles distributed over the cortical surface. In practice the number of vertices on a typical segmented cortical surface ranges from 10000 to 50000. With dipoles having their orientations constrained by the normal to the mesh the number of dipoles corresponds to the number of vertices. When the orientation is not constrained the number of dipoles can be three times bigger.

The conclusion of the latter remarks is that the M/EEG inverse problem with distributed source models is strongly ill posed ($d_m \ll d_x$). The number of unknowns is much bigger than the number of equations. To get estimates of the sources $\mathbf{X}$ given only the the measurements $\mathbf{M}$ requires therefore to consider priors on the solutions.

Setting a prior typically consists in assuming that the solution is small for a given norm denoted just for now by $\| \ \|$. In other words, we assume that a good estimate $\mathbf{X}^*$ of the true source distribution is given by the solution to the following optimization problem:

$$\mathbf{X}^* = \arg\min_{\mathbf{X}} \|\mathbf{M} - \mathbf{G}\mathbf{X}\|_F, \text{ subject to } \|\mathbf{X}\| \leq \eta \ . \tag{3.4}$$

The norm denoted by $\|\|_F$ corresponds to the Frobenius norm of the matrix and the parameter η controls the regularity of the solution. We will refer to $\|\mathbf{M}-\mathbf{G}\mathbf{X}\|_F$ as the *reconstruction error* or the *norm of the residual*. In other words, we want to minimize the reconstruction error while imposing the solution to be small for a given norm.

Problem (3.5) can also be presented as:

$$\mathbf{X}^* = \arg\min_{\mathbf{X}} \|\mathbf{X}\|, \text{ subject to } \|\mathbf{M} - \mathbf{G}\mathbf{X}\|_F \leq \delta \ . \tag{3.5}$$

In this case, we want to minimize the norm of the solution while imposing that the reconstruction error is smaller than δ.

Remark. In statistics, the norm of $\|\mathbf{X}\|$ refers to the *model complexity*. By choosing the norm, i.e, a prior, a model is assumed for the solution. By constraining $\|\mathbf{X}\|$ we impose to the solution to be *simple* according to a given measure of complexity. In other words, the optimal solution is the "simplest" solution, for the model considered, that correctly explains the observed data.

In practice the problem (3.5) is more commonly presented in the Lagrangian formulation:

$$\mathbf{X}^* = \arg\min_{\mathbf{X}} \|\mathbf{M} - \mathbf{G}\mathbf{X}\|_F^2 + \lambda\|\mathbf{X}\|, \lambda > 0 \ . \tag{3.6}$$

The formulations in (3.6), (3.5) and (3.4) are equivalent if the problem is convex. We will refer to (3.6) as the penalized formulation of the problem. The parameter λ controls the "trade-off" between the fidelity to measurements and noise sensitivity. It balances the reconstruction error and the regularity of the solution. The lower the level of noise present in the measurements, the smaller should be the reconstruction error. There is a correspondence between the parameter η and the parameter λ, although this link is usually not explicit.

In the following paragraphs, we will explore a series of approaches that can be cast into this Lagrangian formulation. Focus will be put on a series of variants involving ℓ_2 norms. For priors involving non differentiable constraints and typically ℓ_1 norms, we refer the reader to chapter 4.

3.2.1 The Minimum-Norm solution

In the previous paragraph, we explained that the resolution of the inverse problem with distributed source models leads to an optimization problem where the fit of the data is balanced by a penalization based on a particular norm. In this regard, any distributed inverse solver

is a “minimum-norm” problem. However, in the M/EEG community, the Minimum-Norm solution usually only refers to a minimization of an ℓ_2 norm [101, 220].

Regularization of inverse problems with the ℓ_2 norm was introduced by Tikhonov [203] and is known in statistics as *ridge regression*.

3.2.1.1 Minimum-norm equations

The standard Minimum-Norm solution is obtained by solving:

$$\mathbf{X}^* = \arg\min_{\mathbf{X}} \mathbf{E}(\mathbf{X}) = \arg\min_{\mathbf{X}} \|\mathbf{M} - \mathbf{G}\mathbf{X}\|_F^2 + \lambda\|\mathbf{X}\|_F^2, \lambda > 0 \tag{3.7}$$

The solution of this unconstrained and differentiable problem is obtained by setting the derivative with respect to $\mathbf{X}$ to 0:

$$\begin{aligned} \frac{dE}{d\mathbf{X}} &= 0 \\ \Leftrightarrow -\mathbf{G}^T(\mathbf{M} - \mathbf{G}\mathbf{X}) + \lambda\mathbf{X} &= 0 \\ \Leftrightarrow (\mathbf{G}^T\mathbf{G} + \lambda\mathbf{I})\mathbf{X} &= \mathbf{G}^T\mathbf{M} \\ \Leftrightarrow \mathbf{X} &= (\mathbf{G}^T\mathbf{G} + \lambda\mathbf{I})^{-1}\mathbf{G}^T\mathbf{M} \end{aligned} \tag{3.8}$$

The solution $\mathbf{X}^*$ is given by a simple matrix multiplication:

$$\mathbf{X}^* = (\mathbf{G}^T\mathbf{G} + \lambda\mathbf{I})^{-1}\mathbf{G}^T\mathbf{M} \ . \tag{3.9}$$

The fact that the inverse solution is given by a simple matrix multiplication is a general property of ℓ_2 based methods. This property makes them really attractive, although it can happen that computing the inverse operator is intractable in practice.

To understand this, one can observe that equation (3.9) involves computing the matrix $\mathbf{G}^T\mathbf{G} \in \mathbb{R}^{d_x \times d_x}$, where d_x is the dimensionality of the source space, and inverting a matrix of this size. When considering realistic cortical models this computation becomes impossible.

To give an order of magnitude, a matrix in double precision with 10 000 lines and columns contains 10^8 elements. A double precision number takes 8 bytes in memory which means that the matrix requires $8 \cdot 10^8 = 0.8$ GB of RAM just for storage. On a standard computer, even nowadays, inverting such a matrix can become a computational burden.

To circumvent these limitations, the following trick is used:

Lemma 3.1. *Matrix Inversion (Woodbury matrix identity)*

$$(\mathbf{A} + \mathbf{U}\mathbf{C}\mathbf{V})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{U}(\mathbf{C}^{-1} + \mathbf{V}\mathbf{A}^{-1}\mathbf{U})^{-1}\mathbf{V}\mathbf{A}^1 \tag{3.10}$$

or with $\mathbf{A} = \mathbf{I}$ *and* $\mathbf{C} = \mathbf{I}$

$$(\mathbf{I} + \mathbf{U}\mathbf{V})^{-1} = \mathbf{I} - \mathbf{U}(\mathbf{I} + \mathbf{V}\mathbf{U})^{-1}\mathbf{V} \ . \tag{3.11}$$

Applying equation (3.11) to equation (3.9), with $\lambda = 1$ for simplicity, leads to

$$\begin{aligned} &(\mathbf{G}^T\mathbf{G} + \mathbf{I})^{-1}\mathbf{G}^T \\ &= (\mathbf{I} - \mathbf{G}^T(\mathbf{I} + \mathbf{G}\mathbf{G}^T)^{-1}\mathbf{G})\mathbf{G}^T \\ &= \mathbf{G}^T(\mathbf{I} + \mathbf{G}\mathbf{G}^T)^{-1}(\mathbf{I} + \mathbf{G}\mathbf{G}^T - \mathbf{G}\mathbf{G}^T) \\ &= \mathbf{G}^T(\mathbf{I} + \mathbf{G}\mathbf{G}^T)^{-1} \end{aligned} \tag{3.12}$$

The solution $\mathbf{X}^*$ is now given by:

$$\mathbf{X}^* = \mathbf{G}^T(\mathbf{G}\mathbf{G}^T + \lambda\mathbf{I})^{-1}\mathbf{M} \ , \tag{3.13}$$

which involves the inversion of a small matrix in $\mathbb{R}^{d_m \times d_m}$.

By comparing the MN solution in (3.9) and the LCMV beamformer in (3.1), similarities can be observed. This justifies the limited discussion on beamforming techniques and we refer the reader to [152] where the authors detail how to relate linear beamformers such as LCMV to the Minimum-Norm solutions.

3.2.1.2 Choosing the regularization parameter

The naive but quite efficient approach

The λ is related to the level of noise present in the measurements. Schematically, the larger the noise amplitude, the larger the reconstruction error and the larger λ should be. If λ increases, $\|\mathbf{X}^*\|$ decreases and the reconstruction error increases. In theory, this parameter has to be estimated on each dataset since it depends on the data. However, a strategy exists to get a reasonable estimate of λ. This strategy is used by the Brainstorm toolbox [10].

In order to understand this method, it is necessary to introduce the singular value decomposition (SVD) of $\mathbf{G}$:

$$\mathbf{G} = \mathbf{U}\mathbf{S}\mathbf{V}^T \ ,$$

where the matrices $\mathbf{U}$ and $\mathbf{V}$ are square unitary matrices, *i.e.,* $\mathbf{U}^T\mathbf{U} = \mathbf{I}$ and $\mathbf{V}^T\mathbf{V} = \mathbf{I}$, and the matrix $\mathbf{S}$ is diagonal.

The diagonal entries of $\mathbf{S}$ are the singular values $(s_i)_i$ of $\mathbf{G}$. The $(s_i)_i$ are ordered such that $|s_1| > |s_2| > \cdots > |s_{d_x}|$. By replacing $\mathbf{G}$ by its SVD in equation (3.13) we get:

$$\begin{aligned}
\mathbf{X}^* =& \mathbf{G}^T(\mathbf{U}\mathbf{S}^2\mathbf{U}^T + \lambda\mathbf{I})^{-1}\mathbf{M} \\
=& \mathbf{G}^T(\mathbf{U}(\mathbf{S}^2 + \lambda\mathbf{I})\mathbf{U}^T)^{-1}\mathbf{M} \\
=& \mathbf{G}^T\mathbf{U}(\mathbf{S}^2 + \lambda\mathbf{I})^{-1}\mathbf{U}^T\mathbf{M} \\
=& \mathbf{V}\mathbf{S}(\mathbf{S}^2 + \lambda\mathbf{I})^{-1}\mathbf{U}^T\mathbf{M}
\end{aligned} \tag{3.14}$$

The matrix $(\mathbf{S}^2 + \lambda\mathbf{I})^{-1}$ is also diagonal and its diagonal coefficients are $(s_i^2 + \lambda)_i$. The λ should therefore take a value comparable to the $(s_i^2)_i$. The heuristic choice of λ proposed by the first strategy consists in setting $\lambda = 0.01 s_1^2$. This rule of thumb works quite well in practice.

Brainstorm's implementation also removes singular values for which $s_i^2 < d_m 10^{-7} s_1^2$.

The L-curve

The L-curve approach was originally proposed by Hansen in [103]. The idea is to compute for multiple values of λ the value of the norm $\|\mathbf{X}^*\|$ and the reconstruction error. By plotting the norm of $\|\mathbf{X}^*\|$ as a function of the residual $\|\mathbf{M} - \mathbf{G}\mathbf{X}^*\|$ in *loglog* one gets a curve similar to the curve presented in the illustration figure 3.2. This curves describes an "*L*" and the best λ is obtained at the corner of the curve. It is estimated in practice by looking for the point with the highest curvature. Hansen argues that when λ is smaller than this optimal value, the inverse solver reconstructs part of the noise. In [103], Hansen lists a set of conditions to guarantee that the resulting curve describes an L. One of the conditions is that the signal measured in not too buried in noise. It is observed that when increasing the amplitude of the additive noise, the corner of the curve, used to estimate the λ, becomes harder to see.

When increasing the λ from 0 to ∞, the 2D point $(\|\mathbf{M} - \mathbf{G}\mathbf{X}^*\|, \|\mathbf{X}^*\|)$ goes from the upper left extremity of the curve to the lower right extremity. Thus, for a larger λ the reconstruction error increases.

The generalized cross-validation (GCV)

The generalized cross-validation (GCV) is an alternative to the L-curve from Hansen. It

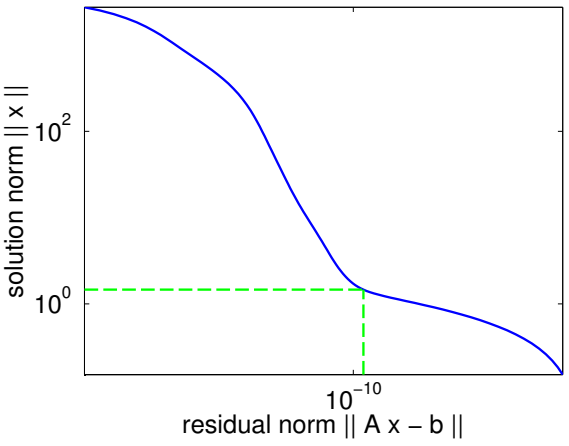


Figure 3.2: L-curve in Minimum-Norm estimator (*loglog* plot). The lambda is estimated by searching the point of highest curvature.

was originally proposed by Wahba and Golub in [90, 218]. The idea behind the GCV is to say that λ is correctly defined if the measurements on $d_m - 1$ sensors can help to predict the measurements of the left out sensor, hence the term cross-validation. Wahba and Golub showed that this prediction error, averaged across all sensors, could be computed with a closed-form for a given λ.

Let us denote the measurements on sensor j by $\mathbf{M}_j$ and by $\mathbf{M}_{|j}$ the measurements obtained by removing the jth sensor. The sources estimated with $\mathbf{M}_{|j}$ are denoted $\mathbf{X}^*_{|j}$ and the leadfield obtained after removing row j is denoted $\mathbf{G}_{|j}$. The jth row of $\mathbf{G}$ is denoted $\mathbf{G}_j$. Using 3.13, we get

$$\mathbf{X}^*_{|j} = \mathbf{G}^T_{|j}(\mathbf{G}_{|j}\mathbf{G}^T_{|j} + \lambda\mathbf{I})^{-1}\mathbf{M}_{|j}$$

The generalized cross-validation error is defined as:

$$\mathcal{G}(\lambda) = \sum_j \|\mathbf{M}_j - \mathbf{G}_j\mathbf{X}^*_{|j}\|^2_F$$

Wahba and Golub have shown that this function of λ can be easily computed with the following formula:

$$\mathcal{G}(\lambda) = \frac{\|\mathbf{M} - \mathbf{G}\mathbf{X}\|^2_F}{(\mathrm{trace}(\mathbf{I} - \mathbf{G}\mathbf{G}^T(\mathbf{G}\mathbf{G}^T + \lambda\mathbf{I})))^2} \tag{3.15}$$

Finding the best λ consists in minimizing this function with respect to λ. Such a function $\mathcal{G}(\lambda)$ is illustrated in figure 3.3.

The equation (3.15) is obtained after assuming that the noise is independent and identically distributed across sensors. When this assumption does not hold the performance of the GCV can be affected. This is a limitation pointed out by Hansen in his presentation of the L-curve method. In practice, this issue can be addressed by pre-whitening the data.

3.2.2 Variants around the minimum-norm solution

We will now present some alternative approaches also based on ℓ_2 penalization, namely, the weighted minimum-norm (WMN), the dSPM [49] and the sLORETA [171] methods. All these methods work time instant by time instant, and therefore do not make use of the temporal correlations of the activations. Finally, we will present an ℓ_2 based method that takes into

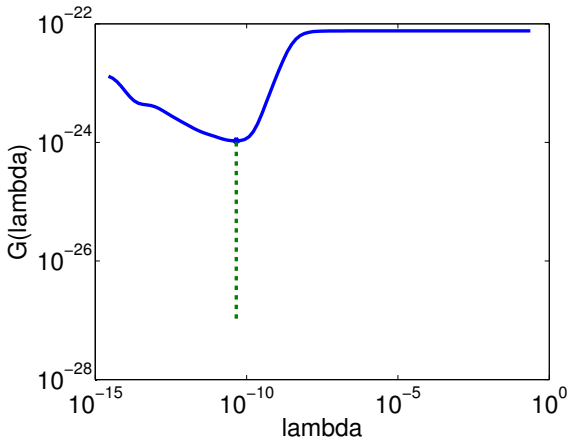


Figure 3.3: Generalized Cross Validation with Minimum-Norm estimator. The vertical line points the minimum of the GCV which provides the value of lambda.

Setting λ with a percentage like exposed in paragraph 3.2.1.2:

```
clear options
% 10 percents corresponds to Brainstorm's default lambda
options.pct = 10;
[X,Ginv] = mn_inverse(M,G,options); % X = Ginv * M;
```

Or manually:

```
clear options
options.lambda = 1e-5;
[X,Ginv] = mn_inverse(M,G,options);
```

Or with generalized cross-validation:

```
clear options
options.use_gcv = true;
[X,Ginv] = mn_inverse(M,G,options);
```

Or with the L-curve:

```
clear options
options.use_lcurve = true;
[X,Ginv] = mn_inverse(M,G,options);
```

Table 3.3: Running a Minimum-Norm with EMBAL .

account this temporal information.

3.2.2.1 The weighted minimum-norm (WMN)

When applying a simple Minimum-Norm solution, each dipolar source is penalized equivalently, although the columns of the leadfield matrix $\mathbf{G}$ are not normalized. Sources that are close to the sensors have a higher forward field, *i.e.,* the effect on the sensors of a small activation is big. As a consequence, the minimum-norm solution is biased towards the superficial sources. The WMN was originally proposed to cope with this problem.

The weighted Minimum-Norm solution corresponds to the problem:

$$\mathbf{X}^* = \arg\min_{\mathbf{X}} \mathcal{E}(\mathbf{X}) = \arg\min_{\mathbf{X}} \|\mathbf{M} - \mathbf{G}\mathbf{X}\|_F^2 + \lambda \|\mathbf{W}\mathbf{X}\|_F^2, \lambda > 0 \ . \tag{3.16}$$

The matrix $\mathbf{W} \in \mathbb{R}^{d_x \times d_x}$ is the weighting matrix. To guarantee a valid penalization, it is necessary to impose that $\mathbf{W}$ is not singular, *i.e.,* $\mathbf{W}^{-1}$ exists. If $\mathbf{W}$ is singular, then the cost function to minimize might not be strictly convex, leading to a non unique solution.

Setting $\mathbf{Y} = \mathbf{W}\mathbf{X}$ and replacing $\mathbf{G}$ by $\tilde{\mathbf{G}} = \mathbf{G}\mathbf{W}^{-1}$ in equation (3.13) leads to

$$\mathbf{Y}^* = \arg\min_{\mathbf{Y}} \mathcal{E}(\mathbf{X}) = \arg\min_{\mathbf{Y}} \|\mathbf{M} - \tilde{\mathbf{G}}\mathbf{Y}\|_F^2 + \lambda \|\mathbf{Y}\|_F^2, \lambda > 0 \ . \tag{3.17}$$

Using (3.13), we get that

$$\mathbf{X}^* = \mathbf{W}^{-1}\mathbf{Y}^* = (\mathbf{W}\mathbf{W}^T)^{-1}\mathbf{G}^T(\mathbf{G}(\mathbf{W}\mathbf{W}^T)^{-1}\mathbf{G}^T + \lambda\mathbf{I})^{-1}\mathbf{M} \ . \tag{3.18}$$

We observe that computing $\mathbf{X}^*$ with a WMN using equation (3.18) requires to be able to invert a big matrix $\mathbf{W}\mathbf{W}^T \in \mathbb{R}^{d_x \times d_x}$, which is what was previously avoided thanks to the matrix inversion lemma.

In order to make the equation (3.18) tractable the weighting matrices $\mathbf{W}$ have to be diagonal, or easily invertible. We will now show how such a diagonal matrix can be chosen in order to avoid the bias towards superficial sources.

The amplitude of the forward field for a dipole close to the sensors is bigger than for a dipole deep in the brain. Hence the standard minimum-norm that penalizes all the dipoles equivalently tends to explain the measurements with superficial dipoles close to the sensors. If a small amplitude for a superficial dipole can explain the measurements, the same effect for a deeper source requires a much bigger amplitude of activation.

Let $\mathbf{G}^i$ denote the ith column of $\mathbf{G}$ and $(w_i)_i$ the diagonal coefficients of $\mathbf{W}$. The norm $\|\mathbf{G}^i\|_2$ is the amplitude of the forward field of the dipole i. By setting $w_i = \|\mathbf{G}^i\|_2^\gamma$ with $\gamma > 0$, the bias is reduced. Typically γ is set to 1 or 0.5. In practice, setting the parameter γ is an issue that has led to its empirical estimation with real and simulated data in [140].

In the case where we do not have access to $(\mathbf{W}\mathbf{W}^T)^{-1}$, the problem has to be solved from the equation obtained after differentiating $\mathcal{E}(\mathbf{X})$:

$$(\mathbf{G}^T\mathbf{G} + \lambda\mathbf{W}^T\mathbf{W})\mathbf{X} = \mathbf{G}^T\mathbf{M}$$

Then for each column of $\mathbf{M}$ the problem can be solved separately using an iterative method such as the conjugate gradient algorithm. Attention should be paid to avoid assembling and storing the matrix $\mathbf{G}^T\mathbf{G}$.

Examples that face this problem are what we call the Laplacian-based penalizations. When trying to impose spatial smoothness on the solution a natural choice for $\mathbf{W}$ is to use the surface Laplacian $\mathbf{L}_{surf}$ of the cortical mesh. This method is generally called in the literature the *maximum smoothness* solution and corresponds to the LORETA solution [172], although LORETA was originally formulated on a grid of dipoles rather than on the cortical mesh.

```
clear options
options.pct = 10;
options.W = sqrt(sum(G.*G))'; % Set weights
X = wmn_inverse(M,G,options);
```

Table 3.4: Running a Weighted Minimum-Norm with EMBAL .

An alternative to LORETA, that also uses a spatial smoothing prior consists in penalizing the estimate with the ℓ_2 norm of the surface gradient ∇_{surf}. The discretization of ∇_{surf} is denoted in matrix form by $\mathbf{D}_{surf}$. This solution is referred to as the HEAT solution [5], since it can be related to the heat equation defined over the mesh. The surface gradient $\mathbf{D}_{surf}$ verifies $\mathbf{D}_{surf}^T\mathbf{D}_{surf} = \mathbf{L}_{surf}$, which implies that the HEAT solution is given by solving

$$(\mathbf{G}^T\mathbf{G} + \lambda\mathbf{L}_{surf})\mathbf{X} = \mathbf{G}^T\mathbf{M} \ .$$

Remark. The gradient and Laplacian are a bit more complex to compute on a tessellated surface than on a grid since it involves a discretization with P1 elements of the activation map. See chapter 2 for more details on P1 elements and their role in finite element methods.

We will come back to the Laplacian-based methods when considering spatiotemporal regularizations.

3.2.2.2 The ℓ_2 priors and Gaussian models

Up to here, ℓ_2 priors have been presented without much attention drawn onto the underlying assumptions. In order to understand the link between ℓ_2 priors and Gaussian models, it is necessary to relate the ℓ_2 penalization model (3.7) to Bayesian estimation.

Let us assume that $\mathbf{M}$ and $\mathbf{X}$ are random variables. According to Bayes' rule, we have:

$$P(\mathbf{X}|\mathbf{M}) = \frac{P(\mathbf{M}|\mathbf{X})P(\mathbf{X})}{P(\mathbf{M})} \ .$$

The optimal $\mathbf{X}$ is obtained by estimating a maximum a posteriori (MAP), *i.e.,* solving:

$$\begin{aligned}\mathbf{X}^*_{MAP} &= \arg\max_{\mathbf{X}} P(\mathbf{X}|\mathbf{M}) \\ &= \arg\max_{\mathbf{X}} P(\mathbf{M}|\mathbf{X})P(\mathbf{X}) \\ &= \arg\min_{\mathbf{X}} -log(P(\mathbf{M}|\mathbf{X})) - log(P(\mathbf{X}))\end{aligned} \tag{3.19}$$

The noise and the source amplitudes are assumed to be Gaussian variables with zero mean and respectively $\mathbf{\Sigma_E}$ and $\mathbf{\Sigma_X}$ as covariances. This leads to:

$$\begin{aligned}\mathbf{X}^*_{MAP} &= \arg\min_{\mathbf{X}} (\mathbf{M}-\mathbf{G}\mathbf{X})^T\mathbf{\Sigma}_\mathbf{E}^{-1}(\mathbf{M}-\mathbf{G}\mathbf{X}) + \mathbf{X}^T\mathbf{\Sigma}_\mathbf{X}^{-1}\mathbf{X} \\ &= \arg\min_{\mathbf{X}} \|\mathbf{M}-\mathbf{G}\mathbf{X}\|^2_{\mathbf{\Sigma_E}} + \|\mathbf{X}\|^2_{\mathbf{\Sigma_X}}\end{aligned} \tag{3.20}$$

We recall that $\|\mathbf{X}\|_{\mathbf{\Sigma}} = \mathrm{trace}(\mathbf{X}^T\mathbf{\Sigma}^{-1}\mathbf{X})$. Using a simple derivation and the matrix inversion lemma 3.1, the solution of this problem is given by:

$$\mathbf{X}^*_{MAP} = \mathbf{\Sigma_X}\mathbf{G}^T(\mathbf{G}\mathbf{\Sigma_X}\mathbf{G}^T + \mathbf{\Sigma_E})^{-1}\mathbf{M} \ . \tag{3.21}$$

We observe that the standard MN corresponds to the case where $\mathbf{\Sigma_X} = \mathbf{I}$ and $\mathbf{\Sigma_E} = \lambda\mathbf{I}$. The Bayesian approach offers a convenient framework for modeling the noise. Note that

in (3.21), the source covariance $\mathbf{\Sigma_X}$ does not need to be inverted. This implies that if the prior is defined on the covariance, there is no need to invert a matrix of size $d_x \times d_x$. Another interesting aspect of the Bayesian approach is its ability to formalize *model selection* methods. This will be detailed later in section 3.3.2.

3.2.2.3 Noise normalized methods: dSPM and sLORETA

The central idea behind noise *normalized methods*, as they are being called in the M/EEG community, is to represent on the cortex not the activity itself, but a *dimensionless* statistical quantity. As we will see, using statistical quantities attenuates the bias towards the superficial sources and also has the advantage of providing a natural way of thresholding the reconstructed estimates.

For simplicity, we will restrict the presentation of both methods to the inverse problem with constrained orientations.

The mathematical foundations of both of these methods come from the knowledge on Student's statistical distributions. Let $(x_i)_i = 1, \ldots, N$ be N normally distributed random samples. We denote by $\overline{x}$ the empirical mean of the $(x_i)_i$ and σ_{emp} the empirical standard deviation:

$$\overline{x} = \frac{1}{N} \sum_{i=1}^{N} x_i \ ,$$

and

$$\sigma_{emp} = \sqrt{\frac{1}{N-1} \sum_{i=1}^{N} (x_i - \overline{x})^2} \ .$$

One can prove that, if the true mean of the $(x_i)_i$ equals μ, then the quantity defined as:

$$\frac{\overline{x} - \mu}{\sigma_{emp}/\sqrt{N}}$$

follows a T-distribution, also called Student's distribution, with $N - 1$ degrees of freedom. This quantity is the T-statistic of the sample.

Definition 3.1 (Student's T-distribution)**.** Student's t-distribution with n degrees of freedom has a probability density function given by:

$$p(x) = \frac{\Gamma(\frac{n+1}{2})}{\sqrt{n\pi}\Gamma(\frac{n}{2})} \left(1 + \frac{x^2}{n}\right)^{-\frac{n+1}{2}} \ ,$$

where Γ is the Gamma function: $\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt$.

When dealing with M/EEG data the classical question is to know whether a given experimental stimulation created an effect in the subject's brain and where. An effect is always measured in comparison with a reference level usually estimated in the period before stimulation, called the baseline period. To take the baseline period into account, a classical preprocessing of the data is called the *baseline correction*. It consists in subtracting to every sensor the mean value of the signal measured during the baseline period. This being done, the model, mentioned is section 3.2.2.2, that assumes that the sources $\mathbf{X}$ and the noise $\mathbf{E}$ have 0 mean holds.

The relevant T-statistic to consider is therefore reduced to:

$$\frac{\overline{x}}{\sigma_{emp}/\sqrt{N}} \ .$$

Without going into more details on statistical tests, under the Gaussianity assumption, if this statistic is large in absolute value, an effect is detected.

The point where dSPM and sLORETA differ is on the way to estimate the empirical standard deviation.

Dynamic Statistical Parametric Mapping (dSPM)

With the dSPM method, the variability of the estimate is assumed to exist due to the additive noise. The sources $\mathbf{X}$ are not supposed to be random variables, unlike the noise $\mathbf{E}$. If we assume that $\mathbf{E}$ is Gaussian and that the covariance matrix of each of its columns is given by $\mathbf{\Sigma_E}$, using the MN inverse formula, we get that the covariance of each of the columns of $\mathbf{X}^*$ is given by $\mathbf{C} = \mathbf{H\Sigma_E H}^T$. The matrix $\mathbf{H}$ denotes here the MN inverse $\mathbf{H} = \mathbf{G}^T(\mathbf{GG}^T + \mathbf{\Sigma_E})^{-1}$ obtained with (3.21). The dSPM method estimates the variances of the sources with the diagonal elements of $\mathbf{C}$. We get:

$$\mathbf{T}_{\mathrm{dSPM}} = \mathbf{RX}^* \ , \tag{3.22}$$

where $\mathbf{R}$ is a diagonal matrix whose coefficients are $\mathbf{R}_{ii} = 1/\sqrt{\mathbf{C}_{ii}}$.

Once the values of $\mathbf{T}$ are computed, we obtain statistical maps, instead of current estimates, and active sources are detected by thresholding them. The threshold can be related to a p-value according to the T-distribution calculated under Gaussianity assumptions.

Since the number of time samples used to calculate the noise covariance matrix C is quite large (typically more than 100), the T-distribution approaches a unit normal distribution. With a large number of samples, the empirical variance approaches the true variance and the T-statistic becomes equivalent to a z-score, a.k.a., a standard score.

sLORETA

In the sLORETA method, the variability is supposed to come also from the sources. If we denote by $\mathbf{\Sigma_X}$ the covariance matrix of the sources, this implies that the covariance of the source estimates is given by $\mathbf{C} = \mathbf{H}(\mathbf{\Sigma_E} + \mathbf{G\Sigma_X G}^T)\mathbf{H}^T$ where the matrix $\mathbf{H}$ is now given by $\mathbf{H} = \mathbf{\Sigma_X G}^T(\mathbf{G\Sigma_X G}^T + \mathbf{\Sigma_E})^{-1}$. The matrix $\mathbf{C}$ is often called the resolution matrix. The result is obtained as with the dSPM method using (3.22).

In practice the source covariance is not observed and without any learning procedure, this covariance has to be fixed a priori. Like for the standard minimum-norm, this covariance is usually set to $\mathbf{\Sigma_X} = \mathbf{I}$.

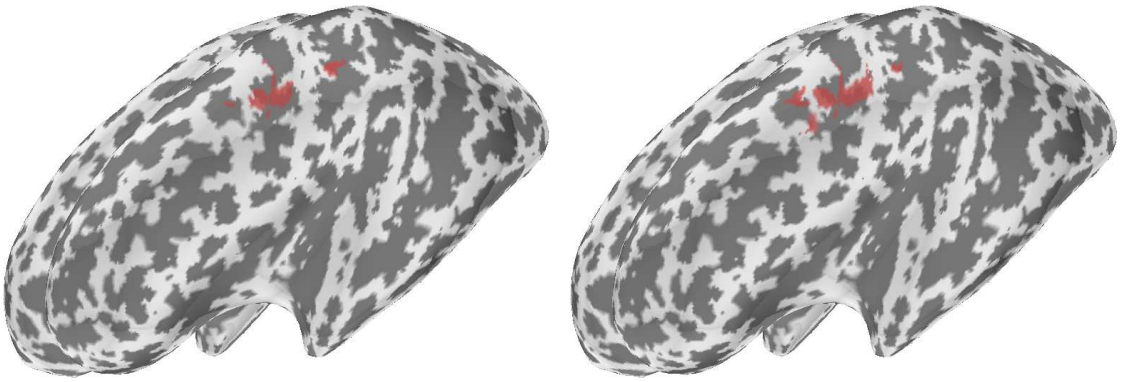
Remark. When considering multiple orientations at each cortical position the statistical quantity considered follows an *F-distribution* (cf. equation 9 in [140]) but the philosophy of the method is the same.

One can prove that in a noiseless case, if only one source is active, the sLORETA method has no localization bias [171].

The dSPM solver is probably more classical, in the sense that the sources are fixed and only the additive noise is a Gaussian random variable. With the sLORETA solver, the sources are also Gaussian random variables, which relates this method to the Bayesian approaches exposed later in this chapter.

3.2.2.4 Spatiotemporal minimum-norm estimation

In the previous paragraphs, we have seen how to set a spatial smoothness prior in the inverse problem. This was done for example with a spatial Laplacian defined over the mesh. We will now see how a similar approach can be used to impose temporal smoothness for the current estimates $\mathbf{X}^*$.



(a) Result obtained with dSPM (b) Result obtained with sLORETA

Figure 3.4: Illustration of thresholded statistical map obtained with the dSPM and sLORETA methods on somato sensory data recorded with MEG.

Let $\mathbf{L}_{time}$ be a temporal Laplacian operator. For a 1D signal $\mathbf{x} = (x_t)_t$, the Laplacian can be approximated by:

$$(L_{time}\mathbf{x})_t = \frac{x_{t-1} - 2x_t + x_{t+1}}{4} \ .$$

With our notations, whereas the spatial Laplacian is applied by multiplication on the left, the temporal Laplacian is applied by multiplication on the right. The currents $\mathbf{X}^*$ are now estimated by solving:

$$\mathbf{X}^* = \arg\min_{\mathbf{X}} \|\mathbf{M} - \mathbf{G}\mathbf{X}\|_F^2 + \lambda\|\mathbf{L}_{surf}\mathbf{X}\|_F^2 + \mu\|\mathbf{X}\mathbf{L}_{time}\|_F^2 \ , \tag{3.23}$$

where $\lambda > 0$ and $\mu > 0$ are the spatial and temporal regularization parameters.

Equation equation (3.23) can be solved very elegantly using Kronecker products. We refer the reader to Appendix A for an introduction to Kronecker products manipulation. We just briefly recall that the stacking operator "*vec*" converts a matrix $\mathbf{A}$ to a vector by stacking all the columns of $\mathbf{A}$. The following equations make an extensive use of the following identity:

$$\text{vec}(\mathbf{AXB}) = (\mathbf{B}^T \otimes \mathbf{A})\text{vec}(\mathbf{X}) \ ,$$

assuming that the dimensions of the matrices $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{X}$ agree.

The cost function in equation (3.23) becomes:

$$\begin{aligned}
&\|\mathbf{M} - \mathbf{G}\mathbf{X}\|_F^2 + \lambda\|\mathbf{L}_{surf}\mathbf{X}\|_F^2 + \mu\|\mathbf{X}\mathbf{L}_{time}\|_F^2 \\
=&\|\text{vec}(\mathbf{M} - \mathbf{G}\mathbf{X})\|_F^2 + \lambda\|\text{vec}(\mathbf{L}_{surf}\mathbf{X})\|_F^2 + \mu\|\text{vec}(\mathbf{X}\mathbf{L}_{time})\|_F^2 \\
=&\|\text{vec}(\mathbf{M}) - (\mathbf{I}\otimes\mathbf{G})\text{vec}(\mathbf{X})\|_F^2 + \lambda\|(\mathbf{I}\otimes\mathbf{L}_{surf})\text{vec}(\mathbf{X})\|_F^2 + \mu\|(\mathbf{L}_{time}^T\otimes\mathbf{I})\text{vec}(\mathbf{X})\|_F^2 \\
=&\|\text{vec}(\mathbf{M}) - \tilde{\mathbf{G}}\text{vec}(\mathbf{X})\|_F^2 + \lambda\|(\mathbf{I}\otimes\mathbf{L}_{surf})\text{vec}(\mathbf{X})\|_F^2 + \mu\|(\mathbf{L}_{time}^T\otimes\mathbf{I})\text{vec}(\mathbf{X})\|_F^2
\end{aligned} \tag{3.24}$$

where $\tilde{\mathbf{G}}$ stands for $\mathbf{I}\otimes\mathbf{G}$.

By differentiating with respect to $\text{vec}(X)$ and setting the derivative to 0 we get:

$$\begin{aligned}
(\tilde{\mathbf{G}}^T\tilde{\mathbf{G}} + \lambda(\mathbf{I}\otimes\mathbf{L}_{surf})^T(\mathbf{I}\otimes\mathbf{L}_{surf}) + \mu(\mathbf{L}_{time}^T\otimes\mathbf{I})^T(\mathbf{L}_{time}^T\otimes\mathbf{I}))\text{vec}(X) &= \tilde{\mathbf{G}}^T\text{vec}(\mathbf{M}) \\
(\tilde{\mathbf{G}}^T\tilde{\mathbf{G}} + \lambda(\mathbf{I}\otimes\mathbf{L}_{surf}^T\mathbf{L}_{surf}) + \mu(\mathbf{L}_{time}\mathbf{L}_{time}^T\otimes\mathbf{I}))\text{vec}(X) &= \tilde{\mathbf{G}}^T\text{vec}(\mathbf{M})
\end{aligned} \tag{3.25}$$

The matrix on the left hand side is extremely large and cannot be stored in memory. However multiplying a vector by this matrix is not very computationally expensive since $\mathbf{L}_{surf}$ and $\mathbf{L}_{time}$ are sparse matrices. The problem is solved using a conjugate gradient method.

As with standard minimum-norm methods, *i.e.,* with no temporal smoothness, the different techniques detailed above can be applied. The parameters can be set by GCV with a function $\mathcal{G}(\lambda, \mu)$. We can imagine an L-curve approach where we would look for a point of highest curvature on a 2D surface parametrized by λ and μ [13, 26]. The dSPM and sLORETA methods can also be extended. The only pitfall here is that, contrary to the standard MN, the inverse matrices cannot be explicitly computed. We need for each pair (λ, μ) to run an iterative solver, which can make the GCV and L-Curve methods particularly time consuming.

3.3 LEARNING-BASED METHODS

In previous sections, the ℓ_2 priors used in the penalization of the inverse problem are defined a priori. Following the explanations in section 3.2.2.2, this means that the proposed methods assume a predefined covariance matrix for the sources. In the following paragraphs, we will present inverse solvers that aim at designing a prior based on the data. The source covariance matrix, *i.e.,* the weights in the ℓ_2 penalization term, is "learned". We will also say that the *model* is learned from the data [201].

For simplicity, we will present the following method in the context of instant-by-instant inverse computation. The methods presented in this section use the Bayesian formulation of the inverse problem with the assumption of no temporal correlations. We recall the Bayesian framework from section 3.2.2.2:

$$p(\mathbf{X}|\mathbf{M}) = \frac{p(\mathbf{M}|\mathbf{X})p(\mathbf{X})}{p(\mathbf{M})} \ . \tag{3.26}$$

where we assume Gaussian variables:

$$\mathbf{E} \sim \mathcal{N}(0, \mathbf{\Sigma_E}) \tag{3.27}$$

$$\mathbf{X} \sim \mathcal{N}(0, \mathbf{\Sigma_X}) \tag{3.28}$$

and an additive model:

$$\mathbf{M} = \mathbf{G}\mathbf{X} + \mathbf{E} \ . \tag{3.29}$$

If $\mathbf{\Sigma_E}$ and $\mathbf{\Sigma_X}$ are known, $\mathbf{X}$ is obtained by maximizing the likelihood which leads to:

$$\mathbf{X}^* = \arg\min_{\mathbf{X}} \|\mathbf{M} - \mathbf{G}\mathbf{X}\|_{\mathbf{\Sigma_E}} + \|\mathbf{X}\|_{\mathbf{\Sigma_X}} \ , \tag{3.30}$$

which leads to:

$$\mathbf{X}^* = \mathbf{\Sigma_X}\mathbf{G}^T(\mathbf{G}\mathbf{\Sigma_X}\mathbf{G}^T + \mathbf{\Sigma_E})^{-1}\mathbf{M} \ .$$

In this framework the prior is an ℓ_2 norm and learning the prior means learning $\mathbf{\Sigma_X}$, *i.e.,* the source covariance matrix. One may also want to learn the noise covariance matrix $\mathbf{\Sigma_E}$. Note that in the WMN framework, learning $\mathbf{\Sigma_X}$ consists in learning the weights.

In the case where $\mathbf{\Sigma_X}$ and $\mathbf{\Sigma_E}$ are not fixed a priori, these parameters define the model commonly denoted $\mathcal{M}$. Bayes' rule can be rewritten:

$$p(\mathbf{X}|\mathbf{M}, \mathcal{M}) = \frac{p(\mathbf{M}|\mathbf{X}, \mathcal{M})p(\mathbf{X}|\mathcal{M})}{p(\mathbf{M}|\mathcal{M})} \ . \tag{3.31}$$

$p(\mathbf{X}|\mathbf{M}, \mathcal{M})$ is called the *posterior*.

$p(\mathbf{M}|\mathbf{X}, \mathcal{M})$ is called the *likelihood*.
$p(\mathbf{X}|\mathcal{M})$ is called the *prior*.
$p(\mathbf{M}|\mathcal{M})$ is called the *model evidence*.

3.3.1 Model selection using a multiresolution approach: MiMS

The Multiresolution Image Model Selection (MiMS) algorithm was proposed by Cottereau *et al*. in [43]. Its principal motivation was to be able to reconstruct spatially extended active regions, but also to quantify their extents. The idea behind this inverse solver is first to estimate weights using a multiresolution approach, and second to use these weights in the WMN framework. The multiresolution step is the learning step. The building blocks of the MiMS source model are parcels of the cortical surface, designed at multiple spatial resolutions in combination with anatomical and functional priors. The sources on the parcels are modeled with current multipoles.

The procedure is iterative and goes from a coarse to a fine resolution. At each iteration k, the source space $\mathcal{M}_k$ is segmented into N_c parcels, or clusters, $(\mathcal{C}_j^k)_j$ of elementary sources:

$$\mathcal{M}_k = \{\mathcal{C}_j^k, j \in [1, N_c]\}. \tag{3.32}$$

A cluster is a cortical region, as illustrated in figure 3.5. The authors call $\mathcal{M}_k$ a *piecewise image model* at resolution k.

From iteration k to iteration $k+1$, some sources are eliminated. After removing these sources, a new *image model* at resolution $k+1$ is defined. The procedure is summarized by the following steps:

1. Design of the piecewise image model $\mathcal{M}_{k+1}$ at resolution $k+1$ from the elementary sources that survived previous eliminatory procedure of Step 3.

2. Compact parametric modeling of regional neural activity from each elementary cluster $\mathcal{C}_j^{k+1}$ in $\mathcal{M}_{k+1}$.

3. Model selection: eliminate using generalized cross-validation (GCV) the least-significant source cluster from $\mathcal{M}_k$ and loop back to Step 1.

We now briefly detail the technical aspects of the modeling and selection steps (see [6, 43] for a full description). At each resolution k, the available cortical sources are clustered in N_c patches $\mathcal{C}_j^k$ of similar surface area. For MEG, Jerbi *et al*. showed in [119] that the current quadrupolar expansion was an adequate model for extended sources (> 5 cm^2) with only 7 moment parameters accounting for the cortical activity generally supported by about 100 elementary dipoles in conventional imaging models (see section 2.3.3). At resolution k, the activity of cluster $\mathcal{C}_j^k$ is modeled up to its quadrupolar expansion about its geometrical centroid (cf. figure 3.5).

Let us denote $\bar{\mathbf{H}}_k$ the forward field of the parcels at resolution k, and $\bar{\mathbf{Q}}_k$ the parameters of the multipoles. The model becomes:

$$\mathbf{M} = \bar{\mathbf{H}}_k \bar{\mathbf{Q}}_k + noise \ . \tag{3.33}$$

The operator $\bar{\mathbf{H}}_k$ has $7N_c$ columns. It is possible to enforce the problem to be overdetermined by assuming that $7N_c$ is smaller than the number of sensors. With an MEG device with 151 sensors, this leads to 21 parcels at each resolution.

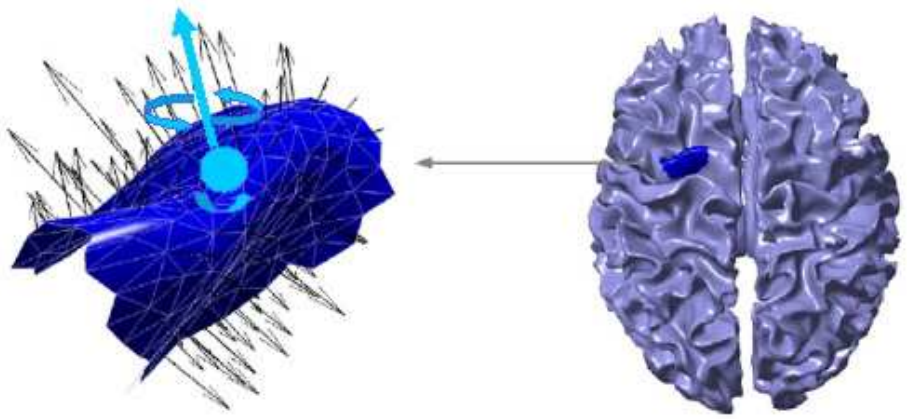


Figure 3.5: Illustration of a $300\ \mathrm{mm}^2$ cortical patch. Moments of current quadrupolar expansion are obtained from the elementary dipole moments supporting the patch geometry, expanded about the parcel centroid.

The least-significant cluster in the source model is removed by computing the Generalized Cross-Validation error (GCV-error) for N_c submodels indexed by j consisting of all clusters in $\mathcal{M}_k$ except $\mathcal{C}_j^k$. The cluster $\mathcal{C}_{j_0}^k$ associated with the smallest GCV-error is supposed to be the least significant and is removed at resolution $k+1$. At step $k+1$, the remaining N_c-1 regions of the cortex are redivided in N_c parcels.

At the end of this exhaustive procedure when no more parcel can be removed, the best model in the GCV-error sense is selected retrospectively (see figure 3.6). As the initial cortical parcellation at $k = 0$ is arbitrary and coarse, the entire process is restarted L times. A weighted summation of the individual GCV errors across the L best models yields a so-called multiresolution clustering (MRC) frequency map of dipole amplitudes. It is this frequency map that is used as weights in a WMN (see (3.18)) in order to get the current estimates.

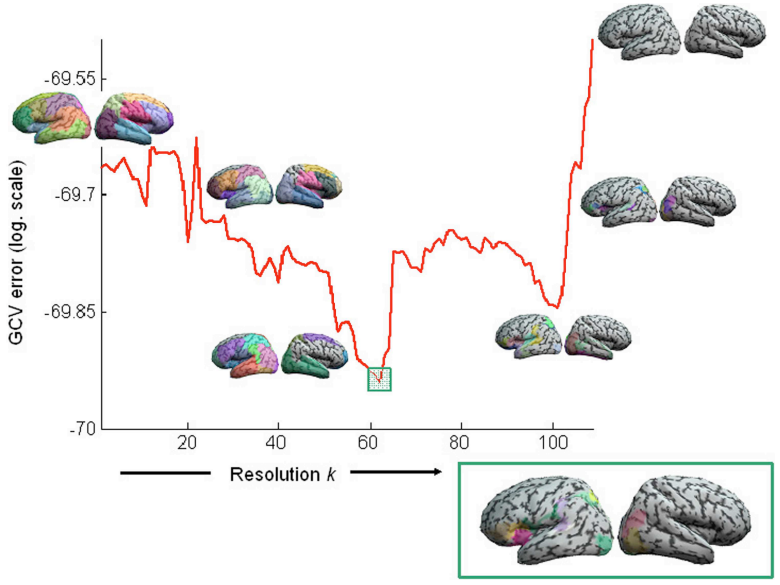


Figure 3.6: GCV error vs. spatial resolution k in semilog scale. A selection of image source models $\mathcal{M}_k$ are shown with their associated cortical parcels. Here, the global minimum of the GCV error is reached for $k^* = 62$ and the corresponding imaging model $\mathcal{M}_k^*$ is magnified in the green boxplot (Adapted from [43]).

This method presents an interesting multiresolution approach and uses a relevant modeling of active parcels via current quadrupoles. By doing so it provides a solution to the challenging problem of estimating the spatial extent of active cortical regions. However, a few critiques remain. First, this approach can only be used with spherical models in MEG (cf. chapter 2 in section 2.2.2). It cannot be used with neither EEG nor MEG when considering non spherical models. Second, it may appear relatively strange to use the frequency map as estimators for source variances. Nevertheless, simulation studies [43] and experimental results on retinotopic mapping [6, 42] obtained with this solver demonstrate its ability to provide good localization results. The retinotopic data are presented in chapter 5.

3.3.2 Restricted Maximum Likelihood (ReML) and Sparse Bayesian Learning (SBL)

More classical Bayesian learning methods use the *evidence framework* in order to learn adaptive parametrized priors from the data itself. A MAP estimate only gives the maximum of the posterior. In practice, this posterior might be multimodal implying that the maximum is not representative of the full posterior. The following approaches aim at estimating the posterior probability mass is order to provide better source estimates. This is done by maximizing the *model evidence*.

Restricted Maximum Likelihood (ReML)

Restricted Maximum Likelihood [105] was introduced to the neuroimaging community by Friston *et al*. [81]. Even if this ReML method has been applied to other problems than M/EEG source imaging, we will present it in this particular context. The general idea of this approach is to estimate hidden variables also called hyperparameters with an iterative procedure that amounts to Expectation-Maximization (EM) update rules [60]. The quantity that drives the learning procedure is the *model evidence*: $p(\mathbf{M}|\mathcal{M})$.

The model is the following:

$$\mathbf{\Sigma_E} = \sum_i \mu_i \mathbf{Q}_\mathbf{E}^i \tag{3.34}$$

$$\mathbf{\Sigma_X} = \sum_i \lambda_i \mathbf{Q}_\mathbf{X}^i \tag{3.35}$$

where the $\mathbf{Q}_\mathbf{E}^i$ and $\mathbf{Q}_\mathbf{X}^i$ are covariance matrices defined a priori. The μ_i and λ_i are positive hyperparameters that need to be learned in order to estimate the "good" model.

Like a standard EM algorithm the procedure consists is maximizing a non-convex likelihood by maximizing a convex surrogate functional for which the optimization is tractable.

The likelihood that needs to be maximized is: $p(\mathbf{M}|\mathcal{M}) = p(\mathbf{M}|\lambda,\mu)$. We have:

$$\begin{aligned}
log(p(\mathbf{M}|\mathcal{M})) &= log\left(\int_{\mathbf{X}} p(\mathbf{M},\mathbf{X}|\mathcal{M})d\mathbf{X}\right) \\
&= log\left(\int_{\mathbf{X}} q(\mathbf{X}|\mathbf{M})\frac{p(\mathbf{M},\mathbf{X}|\mathcal{M})}{q(\mathbf{X}|\mathbf{M})}d\mathbf{X}\right) \\
&\geq \int_{\mathbf{X}} q(\mathbf{X}|\mathbf{M})log\left(\frac{p(\mathbf{M},\mathbf{X}|\mathcal{M})}{q(\mathbf{X}|\mathbf{M})}\right)d\mathbf{X} \quad \text{(Jensen Inequality)} \\
&\geq \left\langle log\left(\frac{p(\mathbf{M},\mathbf{X}|\mathcal{M})}{q(\mathbf{X}|\mathbf{M})}\right)\right\rangle_{q(\mathbf{X}|\mathbf{M})} \\
&\geq \left\langle log\left(\frac{p(\mathbf{M}|\mathbf{X},\mu)p(\mathbf{X}|\lambda)}{q(\mathbf{X}|\mathbf{M})}\right)\right\rangle_{q(\mathbf{X}|\mathbf{M})} \\
&\geq \mathcal{F}(q,\mathcal{M}) = \mathcal{F}(q,\lambda,\mu)
\end{aligned}$$

The surrogate functional $\mathcal{F}$ is called the *free energy*. It is a function of the probability density function q and the hyperparameters (λ,μ).

The EM algorithm alternates an M-step and the E-step.

The M-step finds the Maximum Likelihood (ML) estimate of the hyperparameters:the values of (λ,μ) that maximize $\mathcal{F}$ while keeping q fixed:

$$(\lambda_{k+1},\mu_{k+1}) = \arg\max_{\lambda,\mu} \mathcal{F}(q_k,\lambda,\mu) \ ,$$

where the values of (q,λ,μ) at iteration k are denoted: (q_k,λ_k,μ_k).

Then the E-step consists in updating q using the new values of the hyperparameters $(\lambda_{k+1},\mu_{k+1})$. It is a coordinate ascent on $\mathcal{F}$:

$$q_{k+1} = \arg\max_{q} \mathcal{F}(q,\lambda_{k+1},\mu_{k+1}) \ ,$$

and one can prove that:

$$q_{k+1}(\mathbf{X}|\mathbf{M}) = p(\mathbf{X}|\mathbf{M},\lambda_{k+1},\mu_{k+1}) \ .$$

In their numerous contributions on this subject [78, 79, 80, 81, 146, 177, 178] the authors present the algorithm as a maximization of a Restricted Maximum Likelihood. ReML can be regarded as embedding the E-step into the M-step to provide a single log-likelihood objective function.

The method presented here offers an interesting framework in order to avoid the a priori setting of the ℓ_2 prior, *i.e.,* the source covariance matrix. Its ability to determine the regularisation parameter is also particularly convenient. However, a few critiques can be made about this approach. The optimization method requires a nonlinear search for each M-step (Fisher scoring method) which does not guarantee a positive definite estimated covariance. While shown to be successful in estimating a handful of hyperparameters in [146, 177, 178], this could potentially be problematic when very large numbers of hyperparameters are present. Also, the optimization proposed requires to invert a matrix whose number of rows is equal to the number of hyperparameters. This is another limiting factor when using a large number of priors.

When experimenting with the software provided in the SPM package, we observed that

when using a large number of covariance priors, a fraction of the hyperparameters obtained could be negative-valued. We also observed that the procedure may fail to converge and oscillate between two sets of parameters. It is probably to avoid such problems that the SPM implementation imposes a fixed number of EM iterations. This may appear rather surprising since standard EM algorithms usually require many iterations to converge.

The following approach, based on what is referred to as *Sparse Bayesian Learning (SBL)* and *Automatic Relevance Determination (ARD)*, offers a more principled approach.

Sparse Bayesian Learning and $\gamma - MAP$

The Sparse Bayesian Learning (SBL) approach [142, 160] is an extremely important alternative to the point estimate obtained by simple maximum likelihood. This approach is also based on the maximization of the *model evidence*, but the procedure consists in selecting among a very high number of priors, the ones that fit the best to the data. Contrary to the ReML solver that is limited to a small number of priors, covariance templates in this case, the following method is not. It achieves model selection with a *sparsity* inducing cost function.

In this context, SBL consists in maximizing a tractable Gaussian approximation of the *evidence*, also known as the type-II likelihood or marginal likelihood:

$$p(\mathbf{M}|\mathbf{\Sigma_M}) = \int p(\mathbf{M}|\mathbf{X})p(\mathbf{X}|\mathbf{\Sigma_M})d\mathbf{X} = N(0, \mathbf{\Sigma_M}) \ ,$$

where $\mathbf{\Sigma_M}$ stands for the measurements covariance matrix. This is equivalent to minimizing the negative log marginal likelihood:

$$\begin{aligned} L &= -\log p(\mathbf{M}|\mathbf{\Sigma_M}) \\ &= -\log\left(\frac{1}{\sqrt{(2\pi)^{d_m}\,|\mathbf{\Sigma_M}|}^{d_t}}\exp\left(-\frac{\operatorname{trace}\left(\mathbf{M}^T\mathbf{\Sigma}_M^{-1}\mathbf{M}\right)}{2}\right)\right) . \end{aligned} \tag{3.36}$$

For convenience, the log marginal likelihood is simplified and redefined as:

$$L = d_t \log|\mathbf{\Sigma_M}| + \operatorname{trace}\left(\mathbf{M}^T\mathbf{\Sigma}_\mathbf{M}^{-1}\mathbf{M}\right) \ . \tag{3.37}$$

With the additive model (3.29), $\mathbf{\Sigma_M}$ is given by:

$$\mathbf{\Sigma_M} = \mathbf{G}\mathbf{\Sigma_X}\mathbf{G}^T + \mathbf{\Sigma_E} \ .$$

The model proposed for $\mathbf{\Sigma_X}$ is $\mathbf{\Sigma_X} = \sum_{i=1}^{d_s}\gamma_i\mathbf{C}_i$. The $(\gamma_i)_i$ are the hyperparameters and the $(\mathbf{C}_i)_i$ are the a priori source covariance matrices. With this model the likelihood L is parameterized by the $(\gamma_i)_i$. The $(\mathbf{C}_i)_i$ can be defined as $\mathbf{C}_i = e_i e_i^T$ where e_i is a vector with zeros everywhere except at the ith element, where it is 1. Such covariance matrices model isolated dipolar sources. It is also possible to use general covariance matrices and particularly some that model extended activation, patches. This was proposed in [182].

The cost function (3.37) induces *sparsity* via the term $\log|\mathbf{\Sigma_M}|$, a.k.a., volume-based regularization. It penalizes a measure of the volume formed by the model covariance $\mathbf{\Sigma_M}$. In high dimensions the volume is more efficiently reduced by setting a few dimensions to 0 rather than by diminishing all of them by a small factor. This penalty term promotes a model covariance that is maximally degenerate (or non-spherical), which pushes elements of γ to exactly zero.

If we assume that $\mathbf{C}_i = e_i e_i^T$, the diagonal matrix $\mathbf{\Sigma_X} = \Gamma = \operatorname{diag}(\gamma_i)$ is the prior source covariance matrix which contains the vector of hyperparameters on the diagonal (*i.e.,* the variances). In the ARD framework, the precisions (*i.e.,* inverse variances) are Gamma dis-

tributed. The matrix $\mathbf{\Sigma_E}$ is the noise covariance matrix, which is assumed here to be a multiple of the identity matrix (*e.g.,* $\sigma^2_{\mathbf{E}}\mathbf{I}$, where $\sigma^2_{\mathbf{E}}$ is the noise variance, a hyperparameter that can also be learned from the data or empirically obtained from the measurements).

The optimization is run with an iterative procedure that updates the $(\gamma_i)_i$ at each step. Various update schemes exist to optimize (3.37).

First, the evidence maximization can be achieved by using an Expectation-Maximization update rule:

$$\gamma_i^{(k+1)} = \frac{1}{d_t r_i} \left\| \gamma_i^{(k)} \mathbf{G}_{(:,\,i)}^T \left(\mathbf{\Sigma}_{\mathbf{M}}^{(k)} \right)^{-1} \mathbf{M} \right\|_F^2 + \frac{1}{r_i} \mathrm{trace} \left(\gamma_i^{(k)} \mathbf{I} - \gamma_i^{(k)} \mathbf{G}_{(:,\,i)}^T \left(\mathbf{\Sigma}_{\mathbf{M}}^{(k)} \right)^{-1} \mathbf{G}_{(:,\,i)} \gamma_i^{(k)} \right) \ ,$$

where r_i is the rank of $\mathbf{G}_{(:,\,i)}\mathbf{G}_{(:,\,i)}^T$, and $\mathbf{G}_{(:,\,i)}$ is a matrix grouping the column vectors from $\mathbf{G}$ that are controlled by the same hyperparameter γ_i [227].

It can also be achieved using a fixed-point gradient update rule (called MacKay updates):

$$\gamma_i^{(k+1)} = \frac{\gamma_i^{(k)}}{d_t} \left\| \mathbf{G}_{(:,\,i)}^T \left(\mathbf{\Sigma}_{\mathbf{M}}^{(k)} \right)^{-1} \mathbf{M} \right\|_F^2 \left(\mathrm{trace} \left(\mathbf{G}_{(:,\,i)}^T \left(\mathbf{\Sigma}_{\mathbf{M}}^{(k)} \right)^{-1} \mathbf{G}_{(:,\,i)} \right) \right)^{-1} \ ,$$

or alternatively with:

$$\gamma_i^{(k+1)} = \frac{\gamma_i^{(k)}}{\sqrt{d_t}} \left\| \mathbf{G}_{(:,\,i)}^T \left(\mathbf{\Sigma}_{\mathbf{M}}^{(k)} \right)^{-1} \mathbf{M} \right\|_F \left(\mathrm{trace} \left(\mathbf{G}_{(:,\,i)}^T \left(\mathbf{\Sigma}_{\mathbf{M}}^{(k)} \right)^{-1} \mathbf{G}_{(:,\,i)} \right) \right)^{-\frac{1}{2}} \ .$$

Contrary to the MacKay updates, the latter scheme guarantees that the cost function decreases at each iteration. The proof is based on convex analysis and the introduction of a surrogate function.

With fixed dipole orientations, $\mathbf{G}_{(:,\,i)}$ is a vector, but with unconstrained orientations $\mathbf{G}_{(:,\,i)}$ is a $d_m \times 3$ matrix. For patch source models involving dipoles within a region, $\mathbf{G}_{(:,\,i)}$ is a matrix containing all gain vectors associated with the local patch of cortex. The last two update schemes are much faster than the EM rule (cf. figure 3.7). It can be noticed that one iteration of the gradient-based update is almost identical to the sLORETA algorithm, which is expressed in a completely different framework. Once the optimal hyperparameters have been learned, the source estimates are given by the classical closed-form solution of the Minimum-Norm:

$$\mathbf{X}^* = E[\mathbf{X}|\mathbf{M}; \mathbf{\Sigma_X}] = \mathbf{\Sigma_X} \mathbf{G}^T \left(\mathbf{\Sigma_M} \right)^{-1} \mathbf{M} \ .$$

It is important to note that many SBL algorithms are distinguished by the parameterization of the source covariance matrix. Here the model presented is $\mathbf{\Sigma_X} = \sum_{i=1}^{d_s} \gamma_i \mathbf{C}_i$. It is referred to as the γ-MAP inverse solver [227].

If we did not include the ReML framework from Friston and colleagues in the Sparse Bayesian Learning framework, it is because their approach with a limited number of hyperparameters and diagonal covariances with no non-zero elements, cannot produce sparse estimates.

Comments on Sparse Bayesian Learning. First a remark on the implementation. It can be observed in the update schemes detailed above that once a γ_i is set to 0 it stays at 0. It then requires no more updating. This means that the faster the γ_i are set to 0, the faster the loop over i. It allows for example to run the γ-MAP solver with thousands of covariance templates. Very rapidly only a handful of γ_i are concerned by the update and the algorithm converges even faster.

The second remark is related to the practical use of this solver. With a classical event related experimental setup, a subject is asked to perform a task or simply to respond to a

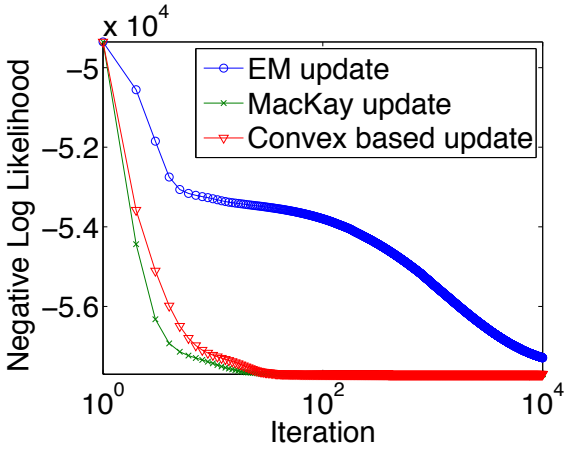


Figure 3.7: Convergence rates observed with the three update schemes (EM, MacKay,Convexity-Based Approach). The EM-based scheme appears clearly as the less efficient. Simulation was run with around 20000 covariance templates.

stimulation multiple times. For each repetition of the experiment, the M/EEG signals are recorded, forming one trial. Let us suppose that d_n trials are recorded and that each trial contains d_t successive time instants. By averaging all the d_n trials, we obtain what is called the evoked response. A first approach consists in using as input successive time frames of the evoked response. The input data could for example be the measured evoked response between 40 and 50 ms or between 20 and 200 ms after the beginning of the stimulation. The problem with the latter example is that source covariance is very likely to change during the time interval. The (γ_i) might be different for early and late brain responses. In order to run the inverse solver on the full temporal data, one might want to consider the possibility of letting γ_i change over time. Indeed, the γ-MAP solver assumes that the noise and prior covariances do not depend on time, which can be a problem for long time interval. Note that this remark is at the origin of our contribution presented in chapter 6.

An alternative that does not suffer from the problem just mentioned consists in using as input the data measured at a given time instant t^* across the d_n repetitions. This procedure provides a set of (γ_i) and localization results at this particular time instant t^* but does not integrate temporal information. Moreover, this approach assumes no variability across trials. The source configuration is considered to be the same at t^* in each repetition. As illustrated in our contribution in chapter 7, this assumption can be questioned.

Finally we would like to mention that our experience with the γ-MAP inverse solver showed that it could provide very accurate results. While performing numerical simulations, it showed its ability to recover very complex source configurations. However, it also proved its sensitivity to the definition of the noise covariance matrix provided as input. A wrong noise covariance matrix can strongly bias the localization result.

With real data, since this solver can provide with focal source estimates, we also observed that the active source could be estimated on the wrong side of a gyrus, a location also confirmed by simple dipole fitting. In this example, the forward modeling was probably at the origin of the problem. The point of this remark is that, a very precise inverse solver, requires a very precise forward modeling.

The γ-MAP inverse solver with $\mathbf{C}_i = e_i e_i^T$ can be used with the code snippet in table 3.5 extracted from EMBAL .

```
clear options
options.noise_cov = C; % Set noise covariance
options.maxit = 500; % Set maximum number of iterations of the gammas
[X,Ginv,gammas] = gmap_inverse(M,G,options);
```

Table 3.5: Running the Gamma-MAP inverse solver with EMBAL .

3.4 CONCLUSION

This chapter was written to provide an overview of the state of the art of M/EEG inverse solvers based on Gaussian assumptions and ℓ_2 priors. The list of solvers presented is fairly long but does not claim to be exhaustive.

All along this chapter, particular attention was drawn to implementation details, mandatory to have software able to deal with realistic datasets. We also took care of providing personal comments on each algorithm by discussing their advantages but also their limitations.

All the algorithms detailed in this chapter (except for MiMS that should soon be integrated in the Brainstorm Toolbox) have been implemented and tested on synthetic and real MEG data. The source code of the solvers and the demo scripts with synthetic and real data are available in a Matlab toolbox called EMBAL (Electro-Magnetic Brain Activity Localization):

`https://gforge.inria.fr/projects/embal`

The following two chapters discuss additional aspects of the distributed source models: priors other than ℓ_2 and frequency domain analyses.

CHAPTER 4

M/EEG INVERSE MODELING WITH NON DIFFERENTIABLE CONSTRAINTS AND SPARSE PRIORS

In chapter 3, focus has been put on distributed inverse solvers with ℓ_2 priors, either with fixed priors or learning-based approaches. When using an ℓ_2 norm to regularize the inverse problem, the cost function to minimize is differentiable and strictly convex, which leads to a convenient solution obtained in closed-form. The ℓ_2 norm is known for its good robustness to noise. However, the ℓ_2 norm suffers from various pitfalls.

The main criticism that is addressed to standard Minimum-Norm solutions, *i.e.,* without Bayesian learning, results from their tendency to smear out the estimated cortical currents, often leading to solutions that are too widely extended. This is intrinsically due to the ℓ_2 norm used to regularize the inverse problem. In order to reduce this effect, a natural choice is to use a regularizing prior that tries to limit the number of active sources, *i.e.,* that introduces *sparsity* in the source space.

At the center of sparse priors are the ℓ_p norms with $p < 2$ and particularly the ℓ_1 norm that achieves sparsity while leading to a convex problem. Such priors lead to non differentiable optimization problems for which numerous optimization methods have been proposed in the last few years [34, 39, 55, 63, 77, 99, 161, 180, 202, 212]. What motivated such an interest is the ability of sparse priors to improve the solution of ill-posed inverse problems present in machine learning and signal processing.

The major motivation for using sparse priors in M/EEG originates from the fact that they provide a natural way to integrate relevant *a priori* information in the inverse problem. Such information includes the number of active sources, the spatial extent of active regions, the spatial or temporal smoothness of the reconstructions or even anatomo-functional knowledge between multiple experimental conditions.

In this chapter, we review previous contributions that introduced sparse priors in the context of M/EEG inverse modeling. The proposed optimization methods are commented on and, eventually, simpler and more efficient algorithms are proposed. We finish this chapter by our contributions that concern the integration of the anatomo-functional knowledge between experimental conditions in the prior [93, 128]. As it will be shown with simulations and real MEG data, this approach offers a principled way to achieve functional mapping with M/EEG with better results than classic MN estimators.

Contents

4.1 WHY USE SPARSE PRIORS?

When using an ℓ_2 norm to regularize the inverse problem, the cost function to minimize is differentiable and strictly convex. The strict convexity implies that the solution necessarily exists and is unique. The differentiability allows to find the optimum by just setting the derivative with respect to $\mathbf{X}$ to 0. This is what provides, in the ℓ_2 case, the solution in closed-form. Here, we focus on strategies that do not exhibit such computational advantages, but that offer the possibility to integrate other prior knowledge on the solution in order to better constrain of the inverse problem.

Historically, sparse priors have been introduced for M/EEG inverse modeling [91, 145] to address a problem raised by standard Minimum-Norm solutions (without Bayesian learning). The MN estimator has a tendency to smear out the estimated cortical currents, often leading to solutions that are too widely extended.

Such behavior of standard MN estimators is intrinsically due to the ℓ_2 norm used to regularize the inverse problem. In order to reduce this effect a natural choice is to use a regularizing prior that tries to limit the number of active sources, *i.e.,* that introduces *sparsity* in the source space. This can be achieved by using ℓ_p (quasi)-norms with $p < 2$.

Definition 4.1 (ℓ_p norms)**.** Let $\mathrm{x} \in \mathbb{R}^I$. The ℓ_p norm for $1 \leq p < \infty$ and the quasi-norm for $0 \leq p < 1$ of the vector x is defined by:

$$\|\mathrm{x}\|_p = \left(\sum_i |x_i|^p \right)^{\frac{1}{p}} \quad . \tag{4.1}$$

Remarks.

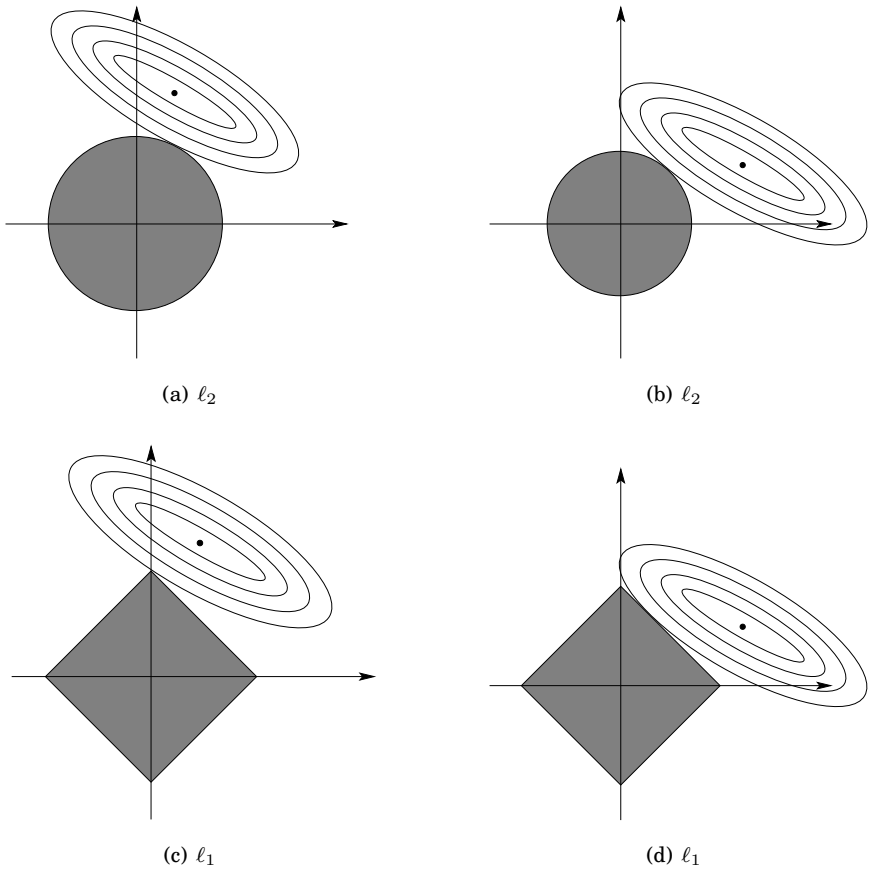
- For $p = 0$, $\|\mathrm{x}\|_0$ is equal to the number of non-zero coefficients of x.
- We only define quasi-norms with $0 \leq p < 1$ since the triangular inequality is not satisfied.

The ℓ_p norms with p close to 1 measure *"diversity"*. It is a notion that is opposed to *"sparsity"*. Hence, minimizing an ℓ_p norm with p close to 1, implies to minimize the diversity and to increase the sparsity. A sparse vector is a vector with a small number of non zero coefficients. Let us illustrate this concept with the classical formulation of the M/EEG inverse problem with distributed source models:

$$\mathbf{X} = \arg\min_{\mathbf{X}} \mathcal{E}(\mathbf{X}) = \arg\min_{\mathbf{X}} \|\mathbf{M} - \mathbf{G}\mathbf{X}\|_F^2 + \lambda\phi(\mathbf{X}), \lambda > 0 \ , \tag{4.2}$$

where $\phi(\mathbf{X}) = \|\mathbf{X}\|_1$ or $\phi(\mathbf{X}) = \|\mathbf{X}\|_2^2$ and $\mathbf{X} \in \mathbb{R}^{d_x \times d_t}$. The ellipses in figure 4.1 represent the isocontours of the datafit, while the circle and the square at the center correspond respectively to the ℓ_2 and ℓ_1 "balls". The isovalues for the ℓ_1 prior are squares while the isovalues for the ℓ_2 prior are circles. At the optimum, the ellipses and the "balls" are tangent. If the tangent point lies over a coordinate axis, a coefficient is set to zero and the solution is sparse. This effect is illustrated in figure 4.1(c).

Popular choices for p are 0 and 1. The ℓ_1 is attractive since it leads to convex optimization problems whereas for $p < 1$ the problems become non convex. When computing the inverse problem instant by instant, the ℓ_1 norm is known in the M/EEG community as the Minimum Current Estimate (MCE) inverse solution [145] while the ℓ_0 norm refers to the FOCUSS inverse solver [91, 92].



(a) ℓ_2 (b) ℓ_2

(c) ℓ_1 (d) ℓ_1

Figure 4.1: Graphical illustration of the difference between ℓ_1 and ℓ_2 norms. The ℓ_1 ball is very likely to be tangent to the isovalues of the datafit represented by the ellipses at its corners as in figure 4.1(c). This produces a solution that sets a coordinate to 0 inducing sparsity. This is not likely to happen with the ℓ_2 norm as illustrated in figure 4.1(a) and figure 4.1(b). However it can happen that both coefficients are non zeros with the ℓ_1 norm as illustrated in figure 4.1(d).

4.2 Inversion with sparse priors: Methods

This section presents different algorithms that can be used for inverse computation with sparse priors. Here, the methods assume that the inverse solution is computed instant by instant like in the original contributions in the M/EEG community [91, 145]. The presentation starts with Iterative Least Squares (IRLS) that consist in iteratively computing WMN solutions with weights updated after each iteration [56, 139]. It is followed by a very brief description of the LARS-LASSO algorithm [63, 202], that is an extremely powerful method for solving the ℓ_1 problem. The LARS-LASSO is a variant of the homotopy method from Osborne [180]. Finally methods based on proximity operators and iterative schemes are detailed [39, 55, 161]. The latter methods are the ones used for the Total Variation (TV) problem and the final contribution on inter-condition sparse priors.

There exist other methods for solving the ℓ_1 penalization problem using iterative thresholding like SPGL1 [212] and fixed point continuation [99] but their limited improvement over simple proximal iterations does not justify their presentation here. The ℓ_1 problem can also be solved with simple coordinate descent [77] or by blockwise coordinate descent also called Block Coordinate Relaxation (BCR) [27]. Depending on the problem of interest, the latter methods can be very competitive.

4.2.1 Iterative Least Squares (IRLS)

IRLS with the ℓ_1 norm

The Minimum Current Estimate (MCE) was introduced in the field of M/EEG by Matsuura and Okabe [145]. As mentioned above, the MCE consists in solving instant by instant, the inverse problem with an ℓ_1 penalization. With $d_t = 1$, the source amplitudes are denoted by a vector $\mathbf{x}$. The inverse problem with an ℓ_1 prior then writes:

$$\mathbf{x}^*_\lambda = \arg\min_{\mathbf{x}} \frac{1}{2}\|\mathbf{m} - \mathbf{G}\mathbf{x}\|^2_F + \lambda\|\mathbf{x}\|_1, \lambda > 0 \ . \tag{4.3}$$

This problem corresponds to the LASSO problem. Originally the authors in [145] proposed to solve this problem using the simplex method by reformulating the problem as a Linear Program (LP). This approach works. However, an optimal solution can be obtained by a relatively simple IRLS algorithm.

The IRLS method is not very competitive in terms of convergence speed compared to cutting edge methods for the ℓ_1 prior and can suffer from numerical instabilities. However, due to its link with Minimum-Norm solutions exposed in previous chapter, we start by presenting this algorithm. It also has the advantage of providing the outline of the FOCUSS algorithm for the ℓ_0 prior.

Let $\mathbf{W}^k$ denote the weighting matrix used in the WMN at iteration k. The matrix $\mathbf{W}^k$ is diagonal, $\mathbf{W}^k = \mathrm{diag}(w_i^k)$. The WMN optimization problem related with $\mathbf{W}^k$ is:

$$\min_{\mathbf{x}} \frac{1}{2}\|\mathbf{m} - \mathbf{G}\mathbf{x}\|^2_2 + \lambda \sum_i w_i^k |x_i|^2, \lambda > 0 \ .$$

In order to give an intuition about the algorithm, it can be noticed that the ℓ_2 norm $\|\mathbf{x}\|_{\mathbf{w},2} = \sum_i w_i^k |x_i|^2$ is equal to the ℓ_1 norm $\|\mathbf{x}\|_1 = \sum_i |x_i|$, when $w_i^k = 1/|x_i|$.

Algorithm 4.1 (IRLS ℓ_1 solver)**.**

- *Initialization:* $\mathbf{W}^0 = \mathbf{I}$

- *Compute:* $\mathbf{x}^{k+1} = (\mathbf{W}^k)^{-1}\mathbf{G}^T(\mathbf{G}(\mathbf{W}^k)^{-1}\mathbf{G}^T + \lambda\mathbf{I})^{-1}\mathbf{m}$
- *Update the weights:* $w_i^{k+1} = 1/|x_i|$
- *Stop if* $\|\mathbf{x}^{k+1} - \mathbf{x}^k\|$ *is smaller than a fixed tolerance value.*

Proposition 4.1. *Algorithm 4.1 converges to a minimizer of* (4.3)*.*

Sketch of the proof. Let $a \in \mathbb{R}_+$. One can prove that:

$$\forall w \in \mathbb{R}_+,\ a \leq f_a(w) = \frac{1}{2}\left(\frac{a^2}{w} + w\right)$$

and that $f_a(a) = a$. The function f_a is strictly convex on $\mathbb{R}_+$.
This gives:

$$\begin{aligned} &\min_{\mathbf{x}} \frac{1}{2\lambda}\|\mathbf{m} - \mathbf{G}\mathbf{x}\|_2^2 + \|\mathbf{x}\|_1 \\ &= \min_{\mathbf{x}} \frac{1}{2\lambda}\|\mathbf{m} - \mathbf{G}\mathbf{x}\|_2^2 + \sum_i |x_i| \\ &= \min_{\mathbf{x},\mathbf{w}} \frac{1}{2\lambda}\|\mathbf{m} - \mathbf{G}\mathbf{x}\|_2^2 + \frac{1}{2}\sum_i \left(\frac{(x_i)^2}{w_i} + w_i\right) \end{aligned} \tag{4.4}$$

The minimization is performed alternatively over $\mathbf{w}$ and $\mathbf{x}$. For fixed $\mathbf{x}$, the $\mathbf{w}$ at optimum is given by: $w_i = |x_i|$. For fixed $\mathbf{w}$, the problem corresponds to a weighted minimum-norm.

This proves that the algorithm 4.1 minimizes at each iteration the energy in (4.3). The reader can refer to [56] for the proof that this iterative scheme actually leads to a minimum. □

It is worth noting that the update rule for $\mathbf{x}$ is equivalent to:

$$\mathbf{x}^{k+1} = \Delta^k\mathbf{G}^T(\mathbf{G}\Delta^k\mathbf{G}^T + \lambda\mathbf{I})^{-1}\mathbf{m}\ ,$$

where Δ^k is the diagonal matrix whose diagonal elements are the $(|x_i^k|)_i$. This prevents division by zero when coefficients vanish as the solution becomes more and more sparse during the iterations. More details on IRLS methods using sparse priors can be found in [56, 139].

As we have seen, the IRLS solver for the LASSO problem (4.3) is extremely simple to implement. However, it may suffer from numerical instabilities due the limited precision of the matrix inversion.

Note that a similar IRLS approach can also be used for a mixed norm involving grouped variables as we will see in the section 4.4.2.

IRLS with the ℓ_0 norm: FOCUSS (FOCal Underdetermined System Solver)

The FOCUSS algorithm as proposed in [91, 92], is an IRLS method used to compute, instant by instant the inverse problem with an ℓ_0 penalization. More generally, it works for ℓ_p norms with $p \leq 1$ [186].

The strategy is very similar to the IRLS solver used to compute the ℓ_1 solution. With the same notations it is given by:

Algorithm 4.2 (IRLS ℓ_0 solver: FOCUSS)**.**

- *Initialization:* $\mathbf{W}^0 = \mathbf{I}$
- *Compute:* $\mathbf{x}^{k+1} = (\mathbf{W}^k)^{-1}\mathbf{G}^T(\mathbf{G}(\mathbf{W}^k)^{-1}\mathbf{G}^T + \lambda\mathbf{I})^{-1}\mathbf{m}$

```
clear options
options.p = 1; % for Lasso
options.maxit = 20; % Set maximum number of iterations
[X] = irls_inverse(M,G,options);
```

Table 4.1: Running an IRLS inverse solver with EMBAL . Here p is set to 1 in order to solve the LASSO problem. If p is set to 0, the FOCUSS solver is used.

- *Update the weights:* $w_i^{k+1} = 1/|x_i|^2$
- *Stop if* $\|\mathbf{x}^{k+1} - \mathbf{x}^k\|$ *is smaller than a fixed tolerance value.*

The difference between algorithm 4.1 and algorithm 4.2 comes from the update rule for the weights. One can prove that updating the weights with $w_i^{k+1} = |x_i|^{p-2}$ leads to a solution of the ℓ_p penalized problem.

We refer to [91, 92] for details on how to circumvent the bias for superficial sources with an ℓ_0 prior.

This IRLS solver can be run with EMBAL using the code snippet in table 4.1.

4.2.2 LARS-LASSO with the ℓ_1 norm

The LARS-LASSO algorithm is a powerful method to solve the LASSO problem (4.3) since it allows to compute the optimum $\mathbf{x}^*$ for all values of λ in one run. The acronym LARS stands for Least angle regression. The LARS-LASSO algorithm is a "path algorithm". It consists in finding the solution for biggest value of λ and following an optimal path of solutions while decreasing the λ. The reason for which such an approach is possible comes from the fact that in the LASSO case this path is piecewise linear.

Let us denote $\mathbf{G}_i$ the ith column of $\mathbf{G}$ and more generally $\mathbf{G}_\Gamma$ the concatenation of the columns of $\mathbf{G}$ whose index belongs to a set of indices Γ.

Proposition 4.2. $\mathbf{x}_\lambda^* = (x_i)_i$ *is optimal iff*

$$\forall i \in 1, \ldots, p, \qquad |{\mathbf{G}_i}^T(\mathbf{m} - \mathbf{G}\mathbf{x}_\lambda^*)| \leq \lambda \tag{4.5}$$

$$\forall i / x_i \neq 0, \quad {\mathbf{G}_i}^T(\mathbf{m} - \mathbf{G}\mathbf{x}_\lambda^*) = \lambda \operatorname{sign}(x_i) \tag{4.6}$$

Sketch of the proof. Let us write the directional derivatives of $\mathcal{E}(\mathbf{x})$ around point $\mathbf{x}$. One can prove that this derivative in direction $\mathbf{u}$ is given by:

$$d_{\mathbf{u}}\mathcal{E}(\mathbf{x}) = -\mathbf{u}^T\mathbf{G}^T(\mathbf{m} - \mathbf{G}\mathbf{x}) + \lambda \sum_i^p \begin{cases} u_i \operatorname{sign}(x_i) & \text{if } x_i \neq 0 \\ |u_i| & \text{if } x_i = 0 \end{cases}$$

In order for $\mathbf{x}$ to be optimal one needs to have, for all $\mathbf{u}$: $d_{\mathbf{u}}\mathcal{E}(\mathbf{x}) > 0$. For a given i, by considering both cases ($x_i \neq 0$ and $x_i = 0$), one gets the constraints in (4.5) and (4.6). □

Let us take Γ as the active set, *i.e.,* $\Gamma = \{i / x_i \neq 0\}$, and ϵ_Γ the sign of the active variables, *i.e.,* $\epsilon_\Gamma = \operatorname{sign}(x_\Gamma)$.

Then equation (4.6) leads to:

$$\mathbf{x}_\Gamma^*(\lambda) = (\mathbf{G}_\Gamma^T\mathbf{G}_\Gamma)^{-1}(\mathbf{G}_\Gamma^T\mathbf{m} - \lambda\epsilon_\Gamma)$$

This equation stays valid as long as the optimality conditions in (4.5) and (4.6) are satisfied. On each interval of λ where they are satisfied, the solution is an affine function of λ. To

```
clear options
options.lambda = 1e-7;
X = lars_inverse(M,G,options);
```

Table 4.2: Running a LASSO inverse solver using the LARS algorithm with EMBAL .

compute the solution, on each interval, one has to follow the direction provided by vector $-(\mathbf{G}_\Gamma^T \mathbf{G}_\Gamma)^{-1}\epsilon_\Gamma$. When an optimality condition does not hold anymore, the active set Γ needs to be updated. We refer the reader to [63] for more details on the update rules. As a result, we obtain the solution for all possible λ.

After each update, the optimal direction is obtained by inverting a square matrix in $\mathbb{R}^{d_\Gamma \times d_\Gamma}$, where d_Γ stands for the size of the active set. The complexity of the LARS is cubic in the number of active variables and can therefore be outperformed by other methods when the optimal active set is big.

In practice, the inverse can be efficiently computed using prior inversions. Between two calls of the inverse method, a variable can either be added to the active set or removed. Therefore one line and one column are being added or removed to the matrix to invert. In this case, it is possible to update the inverse, using tricks like for example the matrix inversion lemma or efficient Cholesky updates, yielding significant speed ups.

This LARS algorithm can be run with EMBAL using the code snippet in table 4.2.

4.2.3 Proximity operators and iterative schemes

When $\mathbf{G} = \mathbf{I}$, *i.e.,* there is no smoothing kernel or "convolution" operator, the problem in (4.3) corresponds to:

$$\begin{aligned} \mathbf{x}^* &= \arg\min_{\mathbf{x}} \frac{1}{2}\|\mathbf{y} - \mathbf{x}\|_2^2 + \lambda\|\mathbf{x}\|_1, \lambda > 0 \\ &= \arg\min_{\mathbf{x}} \sum_i \left(\frac{1}{2}(y_i - x_i)^2 + \lambda|x_i|\right) \end{aligned} \tag{4.7}$$

In this case, the problem can be solved coordinate by coordinate, and one can easily prove that an exact solution is given by a soft thresholding [61]:

$$\forall i, x_i^* = y_i \left(1 - \frac{\lambda}{|y_i|}\right)^+ \quad , \tag{4.8}$$

where, by definition, we have $(x)^+ \stackrel{\text{def}}{=} \max(x, 0)$. By convention, $\cdot/0 = 0$, meaning that if $y_i = 0$ then $x_i^* = 0$.

While the problem can be solved analytically when there is no convolution operator, $\mathbf{G} = \mathbf{I}$, it is not the case with a general matrix $\mathbf{G}$. In order to solve the general case in (4.3), one needs to introduce the notion of *proximity operator*, well known in convex analysis, and the iterative *forward-backward* algorithm [120].

Definition 4.2 (Proximity operator)**.** Let $\phi : \mathbb{R}^P \to \mathbb{R}$ be a lower semicontinuous, convex function. The proximity operator associated with ϕ and $\lambda \in \mathbb{R}_+$ denoted by $\text{prox}_{\lambda\phi} : \mathbb{R}^P \to \mathbb{R}^P$ is given by

$$\text{prox}_{\lambda\phi}(\mathbf{y}) = \arg\min_{\mathbf{x}\in\mathbb{R}^P} \frac{1}{2}\|\mathbf{y} - \mathbf{x}\|_2^2 + \lambda\phi(\mathbf{x}) \ .$$

Remark. When ϕ is the ℓ_1 norm, the proximity operator is given by a soft thresholding (4.8).

The following algorithm provides an optimization strategy for:

$$\mathbf{x}_\lambda^* = \arg\min_{\mathbf{x}} \frac{1}{2}\|\mathbf{m} - \mathbf{G}\mathbf{x}\|_2^2 + \lambda\phi(\mathbf{x}), \lambda > 0 \ . \tag{4.9}$$

where ϕ is a lower semicontinuous, convex function.

Algorithm 4.3 (Forward-Backward Proximal iterations)**.**

- Initialize*: Choose* $\mathbf{x}^{(0)} \in \mathbb{R}^{d_x}$ *(for example* $\mathbf{0}$*).*
- Iterate*:*
$$\mathbf{x}^{(k+1)} = \text{prox}_{\mu\lambda\phi}\left(\mathbf{x}^{(k)} + \mu\mathbf{G}^T(\mathbf{m} - \mathbf{G}\mathbf{x}^{(k)})\right)$$
where $0 < \mu < 2|||\mathbf{G}^T\mathbf{G}|||^{-1}$.
- Stop *if* $\|\mathbf{x}^{(k+1)} - \mathbf{x}^{(k)}\|/\|\mathbf{x}^{(k)}\|$ *is smaller than a fixed tolerance criterion.*

Theorem 4.3. *Algorithm 4.3 converges to a minimizer of* (4.9)*, for any choice of* $\mu \in [\epsilon, 2|||\mathbf{G}^T\mathbf{G}|||^{-1} - \epsilon]$ *,* $\epsilon > 0$*.*

Proof. The convergence of this algorithm is guaranteed by results by Combettes *et al*. in [39] using the properties of forward-backward proximal iterations originally proposed by Moreau in [120]. Daubechies *et al*. in [55] prove a similar result but end up with the condition $0 < \mu < |||\mathbf{G}^T\mathbf{G}|||^{-1}$. □

In practice, we set $\mu = |||\mathbf{G}^T\mathbf{G}|||^{-1}$ as it appears to provide better results.

Remarks.

- The stopping criterion proposed here is based on the ratio $\|\mathbf{x}^{(k+1)} - \mathbf{x}^{(k)}\|/\|\mathbf{x}^{(k)}\|$. This is certainly not the most principled way to stop the algorithm but it appears in our context to provide an acceptable strategy. A more rigorous criteria could be based on the size of the *duality gap* [20]. This is, however, may not be trivial to compute for certain priors.
- The iterations in algorithm 4.3 are also called *Landweber* iterations.

The solution of (4.3) is obtained by setting $\phi = \|\ \|_1$ and using for $\text{prox}_{\mu\lambda\|\ \|_1}$ the soft thresholding detailed in (4.8). This algorithm is called ISTA (Iterative Soft Thresholding Algorithm) in the signal processing community.

In order to better understand the idea behind this method, one needs to notice that the term $\mathbf{G}^T(\mathbf{m} - \mathbf{G}\mathbf{x}^{(k)})$ corresponds to the gradient of the reconstruction error. The algorithm can therefore be understood as an alternated minimization over the regularization term, with the proximity operator, and over the reconstruction error via simple gradient descent.

The convergence speed of this method can however be quite slow, especially if the conditioning of the matrix $\mathbf{G}$ is bad. This is mainly due to the fact that the step size scaled by μ is fixed and can be relatively small. This issue can be fixed using more complex optimization schemes proposed by Nesterov[161]. A convenient rewriting of these algorithms using proximity operators is presented in [224] in chapter 4. We rewrite it here, with our notation:

Algorithm 4.4 (Nesterov scheme with proximity operators)**.**

- Initialize*: Choose* $\mathbf{x}^{(0)} \in \mathbb{R}^{d_x}$ *(for example* $\mathbf{0}$*).*
- Set auxiliary variables*:* $a = 0$*,* $\mathbf{g} = 0$*,* $\mu = |||\mathbf{G}^T\mathbf{G}|||^{-1}$.
- Iterate*:*

- $t = 2\mu$
- $b = \frac{t+\sqrt{t^2+4ta}}{2}$
- $\mathbf{v} = \text{prox}_{a\lambda\phi}(\mathbf{x}^{(0)} - \mathbf{g})$
- $\mathbf{u} = \frac{a\mathbf{x}^{(k)}+b\mathbf{v}}{a+b}$
- $\mathbf{x}^{(k+1)} = \text{prox}_{\lambda\mu\phi}(\mathbf{u} + \mu\mathbf{G}^T(\mathbf{m} - \mathbf{G}\mathbf{x}^{(k)}))$
- $\mathbf{g} = \mathbf{g} - b\mathbf{G}^T(\mathbf{m} - \mathbf{G}\mathbf{x}^{(k+1)})$
- $a = a + b$

- Stop *if* $\|\mathbf{x}^{(k+1)} - \mathbf{x}^{(k)}\|/\|\mathbf{x}^{(k)}\|$ *is smaller than a fixed tolerance criterion.*

Nesterov proved that $\mathcal{E}(\mathbf{x}^{(k)}) - \mathcal{E}(x^*)$ decreases in $O(1/k^2)$ and that this is the best convergence rate that can be achieved by a first-order method. A first-order method is a method that only requires to compute gradients, *i.e.,* first derivatives. What can however be reproached to this algorithm is that when modifying λ, the "history" of the gradients that have been used in this multistep approach needs to be cleared. In this sense, knowing $\mathbf{x}^*$ for a given λ does not help much to find the optimum for another λ even close. Nesterov's scheme does not completely benefit from "warm restarts".

In practice, for numerical stability of the algorithm, we observed that increasing μ can be necessary. In our implementation we set $\mu = (1.05 \cdot |||\mathbf{G}^T\mathbf{G}|||)^{-1}$.

Each iteration in algorithm 4.4 contains two gradient computations and two calls to the proximity operator. It is twice more than an iteration in algorithm 4.3 . However the cost is fully justified by the speed of convergence (cf. figure 4.2).

Remark. We presented Nesterov's optimization scheme using a quadratic reconstruction error. However Nesterov's scheme can be applied to any functional of the form:

$$\mathcal{J}(\mathbf{x}) = \psi(\mathbf{x}) + \phi(\mathbf{x}) \ ,$$

where ψ is differentiable with an L-Lipschitz derivative and ϕ a lower semicontinuous convex function. In our case the Lipschitz constant L is given by the spectral norm $|||\mathbf{G}^T\mathbf{G}|||$.

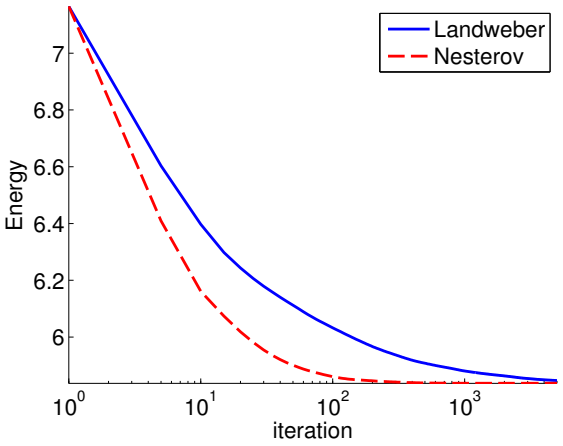


Figure 4.2: Comparison of convergence speed between Landweber and Nesterov iterative schemes. Computation was run with a real MEG leadfield and an ℓ_1 prior.

```
clear options
options.maxit = 500;
options.mode = 'nesterov';
% options.mode = 'landweber'; % Or use landweber
options.lambda = 1e-7;
options.penalty = 'l1'; % Use an L1 prior \ie a Lasso
X = prox_inverse(M,G,options);
```

Table 4.3: Running an inverse solver using proximity operators with EMBAL .

Optimizing with a constraint on the reconstruction error

In practice, it happens that the inverse problem presented in its penalized form as in (4.9) is not the most natural way to constrain the inverse problem. One may have a good estimate of the noise amplitude, with the baseline period for example, and therefore of the norm of the residual, a.k.a., the reconstruction error. Hence a natural formulation of the constrained problem is:

$$\mathbf{x}^*_\lambda = \arg\min_{\mathbf{x}} \phi(\mathbf{x}), \text{s.t.} \|\mathbf{m} - \mathbf{G}\mathbf{x}\|_2 \leq \delta, \delta > 0 \ . \tag{4.10}$$

In order to solve this problem, we propose the following empirical strategy that consists in updating the λ after p forward-backward iterations using the current value of the reconstruction error. The parameter λ now depends on the iteration number and is indexed by k: $\lambda^{(k)}$.

Algorithm 4.5 (Forward-Backward Proximal iterations with constraint on the residual)**.**

- Initialize*: Choose* $\mathbf{x}^{(0)} \in \mathbb{R}^{d_x}$ *(for example* $\mathbf{0}$*).*
- Iterate*:*
$$\mathbf{x}^{(k+1)} = \text{prox}_{\mu\lambda^{(k)}\phi}\left(\mathbf{x}^{(k)} + \mu\mathbf{G}^T(\mathbf{m} - \mathbf{G}\mathbf{x}^{(k)})\right)$$
where $0 < \mu < |||\mathbf{G}^T\mathbf{G}|||^{-1}$.
- Update λ*: If* $t + 1 \equiv 0$ *(mod* p*)*
$$\lambda^{(k+1)} = \lambda^{(k)} \frac{\delta}{\|\mathbf{m} - \mathbf{G}\mathbf{x}^{(k)}\|_2}$$
- Stop *if* $\|\mathbf{x}^{(k+1)} - \mathbf{x}^{(k)}\|/\|\mathbf{x}^{(k)}\|$ *is smaller than a fixed tolerance criterion.*

In practice, updating λ every 10 iterations, *i.e.,* $p = 10$, is a good trade-off between the computational cost of computing the residual and the speed of the convergence observed.

This trick that consists in dynamically changing the λ was proposed by Chambolle in [30] when regularizing with the Total Variation. Our experience confirms that it also works with all the convex priors detailed in this chapter.

It is possible to run this solver with a constraint on the reconstruction error with EMBAL using the code snippet provided in table 4.4.

Mixing sparse-priors and ℓ_2 priors

A sparse prior like a ℓ_1 norm leads to a convex problem. In order to guarantee the uniqueness of the solution, one needs a strictly convex problem. An easy way to achieve this is to

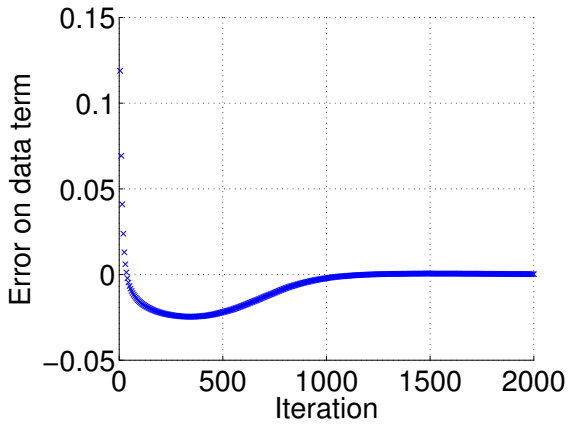


Figure 4.3: Convergence of the optimization with constraint on the reconstruction error. The error corresponds to $\delta - \|\mathbf{m} - \mathbf{G}\mathbf{x}\|_2$ and converges to 0.

```
clear options
options.maxit = 3000;
options.delta = 1e-2;
options.mode = 'landweber';
options.penalty = 'l1'; % Use an L1 prior \ie a Lasso
X = prox_inverse(M,G,options);
```

Table 4.4: Running an inverse solver using proximity operators with EMBAL and a constraint on the reconstruction error.

add an ℓ_2 term to the cost function to minimize:

$$\mathbf{x}^*_\lambda = \arg\min_{\mathbf{x}} \frac{1}{2}\|\mathbf{m} - \mathbf{G}\mathbf{x}\|_2^2 + \lambda \left((1-\rho)\phi(\mathbf{x}) + \frac{1}{2}\rho\|\mathbf{L}\mathbf{x}\|_2^2 \right), \lambda > 0 \ . \tag{4.11}$$

Adding an ℓ_2 term to the Lasso problem was proposed in [239] and is called in the literature Elastic-Net. Using a gradient for the operator $\mathbf{L}$ was introduced under the name of the *Smooth-Lasso* in [106]. In practice adding an ℓ_2 term to the LASSO problem tends to produce results that are less sensitive to the noise inherent to any real dataset. When using a gradient operator for $\mathbf{L}$ it also promotes neighboring active dipoles, which produces in the context of M/EEG inverse modeling spatially consistent active patterns. This idea of mixing priors can be found in [210].

Solving (4.11) can be done elegantly by noticing that the functional can be rewritten:

$$\mathbf{x}^*_\lambda = \arg\min_{\mathbf{x}} \frac{1}{2}\|\mathbf{m}' - \mathbf{G}'\mathbf{x}\|_2^2 + \lambda(1-\rho)\phi(\mathbf{x}), \lambda > 0 \ . \tag{4.12}$$

where:

$$\mathbf{m}' = \begin{pmatrix} \mathbf{m} \\ 0 \end{pmatrix}$$

and

$$\mathbf{G}' = \begin{pmatrix} \mathbf{G} \\ \sqrt{\lambda\rho}\mathbf{L} \end{pmatrix} \ .$$

The algorithms 4.3 and 4.4 can now be reformulated for the new problem (4.11).

```
clear options
options.maxit = 3000;
options.mode = 'nesterov';
options.rho = 0.1;
options.L = mesh_gradientP1(points,faces);
options.lambda = 1e-7;
options.penalty = 'l1'; % Use an L1 prior \ie a Lasso
X = prox_inverse(M,G,options);
```

Table 4.5: Running an inverse solver with two priors (one non differentiable and an ℓ_2 term) using proximity operators with EMBAL .

Algorithm 4.6 (Forward-Backward Proximal iterations with an additive ℓ_2 prior)**.**

- Initialize*: Choose* $\mathbf{x}^{(0)} \in \mathbb{R}^{d_x}$ *(for example* $\mathbf{0}$*).*
- Iterate*:*

$$\mathbf{x}^{(k+1)} = \text{prox}_{\mu\lambda(1-\rho)\phi}\left(\mathbf{x}^{(k)} + \mu((1-\rho)\mathbf{G}^T(\mathbf{m} - \mathbf{G}\mathbf{x}^{(k)}) - \rho\mathbf{L}^T\mathbf{L}\mathbf{x})\right)$$

  where $0 < \mu < 2/(|||\mathbf{G}^T\mathbf{G}||| + \lambda\rho|||\mathbf{L}^T\mathbf{L}|||)$.

- Stop *if* $\|\mathbf{x}^{(k+1)} - \mathbf{x}^{(k)}\|/\|\mathbf{x}^{(k)}\|$ *is smaller than a fixed tolerance criterion.*

Algorithm 4.7 (Nesterov scheme with an additive ℓ_2 prior)**.**

- Initialize*: Choose* $\mathbf{x}^{(0)} \in \mathbb{R}^{d_x}$ *(for example* $\mathbf{0}$*).*
- Set auxiliary variables*:* $a = 0$*,* $\mathbf{g} = 0$*,* $\mu = (|||\mathbf{G}^T\mathbf{G}||| + \lambda\rho|||\mathbf{L}^T\mathbf{L}|||)^{-1}$.
- Iterate*:*
  - $t = 2\mu$
  - $b = \frac{t+\sqrt{t^2+4ta}}{2}$
  - $\mathbf{v} = \text{prox}_{a\lambda(1-\rho)\phi}(\mathbf{x}^{(0)} - \mathbf{g})$
  - $\mathbf{u} = \frac{a\mathbf{x}^{(k)}+b\mathbf{v}}{a+b}$
  - $\mathbf{x}^{(k+1)} = \text{prox}_{\mu\lambda(1-\rho)\phi}(\mathbf{u} + \mu(\mathbf{G}^T(\mathbf{m} - \mathbf{G}\mathbf{x}^{(k)}) - \lambda\rho\mathbf{L}^T\mathbf{L}\mathbf{x}^{(k)}))$
  - $\mathbf{g} = \mathbf{g} - b\mathbf{G}^T(\mathbf{m} - \mathbf{G}\mathbf{x}^{(k+1)}) + \lambda\rho\mathbf{L}^T\mathbf{L}\mathbf{x}^{(k+1)}$
  - $a = a + b$
- Stop *if* $\|\mathbf{x}^{(k+1)} - \mathbf{x}^{(k)}\|/\|\mathbf{x}^{(k)}\|$ *is smaller than a fixed tolerance criterion.*

This algorithm can run with EMBAL using the code in table 4.5.

4.3 Sparsity and spatially extended activations: The Total Variation

The ℓ_1 and ℓ_0 priors are not adapted to spatially extended activations. In order to understand this, let us consider the case where the SNR is particularly bad. The lower is the SNR, the bigger is the regularization parameter λ. By increasing λ, the sparsity of $\mathbf{x}^*$ is increased. This implies that the number of active dipoles gets smaller, *i.e.,* the extent of the active region becomes more and more limited. Active regions become focal. Therefore, by increasing

the additive noise at the sensor level, the ℓ_1 and ℓ_0 priors create a bias for very focal source distributions. To tackle this limitation, the Total Variation (TV) prior can be used.

The TV prior penalizes the solution with the ℓ_1 norm of the gradient, in the present case the surface gradient: $\mathrm{TV}(\mathbf{x}) = \|\nabla_{surf}\mathbf{x}\|_1$. It is a seminorm, since it can be equal to 0 even if $\mathbf{x}$ is non-zero. Assuming $\mathbf{x}$ corresponds to a discretization with P1 elements on the tessellation, $\nabla_{surf}\mathbf{x}$ is a constant vector of $\mathbb{R}^3$ on each triangle. Let us denote it $(\nabla_p^x\mathbf{x}, \nabla_p^y\mathbf{x}, \nabla_p^z\mathbf{x}) \in \mathbb{R}^3$, where p indexes triangles. More details on how to compute the surface gradient with a P1 discretization can be found in [4]. With these notations, $\mathrm{TV}(\mathbf{x})$ can be written:

$$\mathrm{TV}(\mathbf{x}) = \sum_p \sqrt{(\nabla_p^x\mathbf{x})^2 + (\nabla_p^y\mathbf{x})^2 + (\nabla_p^z\mathbf{x})^2}$$

It can be proved that $\mathrm{TV}(\mathbf{x})$ is equal to the sum of the lengths of the isolevels of $\mathbf{x}$ [143]:

$$\mathrm{TV}(\mathbf{x}) = \int_{\mathbb{R}} \mathrm{length}(C_t)dt \ ,$$

where $C_t = \{r \in \text{mesh s.t. } \mathbf{x}(r) = t\}$ is the isolevel, or levelset, at level t. The consequence is that penalizing the inverse problem with the Total Variation tends to produce piecewise constant reconstructions with regular borders.

The TV norm has been widely used for image deconvolution and restoration [12, 30, 31]. It is related in functional analysis to the space of functions with bounded variations. For the M/EEG inverse problem the TV prior was originally proposed in [5]. Here, we argue that the TV regularization can be casted into the general framework of M/EEG inverse solvers with sparsity inducing priors and that the iterative optimization schemes detailed above for the standard ℓ_1 norm can be directly adapted to invert with a TV prior.

The optimization problem considered is given by:

$$\mathbf{x}^* = \arg\min_{\mathbf{x}} \frac{1}{2}\|\mathbf{m} - \mathbf{G}\mathbf{x}\|_F^2 + \lambda TV(\mathbf{x}), \lambda > 0 \ . \tag{4.13}$$

The TV is based on a ℓ_1 norm, it is convex and lower semicontinuous. Hence, assuming one knows how to compute the proximity operator associated to the TV norm, the Forward-Backward iterations and Nesterov schemes can be used to optimize (4.13).

The proximity operator $\mathrm{prox}_{\lambda\| \ \|_{TV}}$ corresponds to following problem:

$$\mathbf{x}^* = \arg\min_{\mathbf{x}} \frac{1}{2}\|\mathbf{y} - \mathbf{x}\|_F^2 + \lambda TV(\mathbf{x}), \lambda > 0 \tag{4.14}$$

This problem is known in the literature as the ROF (Rudin, Osher and Fatemi) problem [189]. Various solvers have been proposed in the literature to solve this problem [31, 32]. Here, we detail the dual approach from Chambolle [30]. It consists in using a gradient-based algorithm for solving the dual problem that interestingly is a smooth optimization problem over a convex set.

The duality between the ℓ_1 norm and the ℓ_∞ norm reads

$$\begin{aligned} TV(\mathbf{x}) &= \|\nabla\mathbf{x}\|_1 \\ &= \max_{\|\mathbf{z}\|_\infty \leq 1} \langle \nabla\mathbf{x}, \mathbf{z}\rangle \ . \end{aligned} \tag{4.15}$$

We also need the adjoint relation between the gradient and the divergence operator:

$$\langle \nabla\mathbf{x}, \mathbf{y}\rangle = -\langle \mathbf{x}, \mathrm{div}\,\mathbf{y}\rangle \tag{4.16}$$

Minimization in equation (4.14) becomes:

$$\begin{aligned}
&\min_{\mathbf{x}} \left(\frac{1}{2}\|\mathbf{y}-\mathbf{x}\|_2^2 + \lambda TV(\mathbf{x}) \right) \\
=&\lambda \min_{\mathbf{x}} \left(\frac{1}{2\lambda}\|\mathbf{y}-\mathbf{x}\|_2^2 + \max_{\|\mathbf{z}\|_\infty \leq 1} \langle \nabla \mathbf{x}, \mathbf{z} \rangle \right) \\
=&\lambda \max_{\|\mathbf{z}\|_\infty \leq 1} \left(\min_{\mathbf{x}} \left(\frac{1}{2\lambda}\|\mathbf{y}-\mathbf{x}\|_2^2 + \langle \nabla \mathbf{x}, \mathbf{z} \rangle \right) \right) \\
=&\lambda \max_{\|\mathbf{z}\|_\infty \leq 1} \left(\min_{\mathbf{x}} \left(\frac{1}{2\lambda}\|\mathbf{y}-\mathbf{x}\|_2^2 - \langle \mathbf{x}, \operatorname{div} \mathbf{z} \rangle \right) \right)
\end{aligned} \tag{4.17}$$

The computation of the minimum and the maximum above can be exchanged because the optimization over $\mathbf{x}$ is convex and the optimization over $\mathbf{z}$ is concave (see for example [188]).

By setting the derivative with respect to $\mathbf{x}$ to zero one gets:

$$\mathbf{x}^* = \mathbf{y} + \lambda \operatorname{div} \mathbf{z}$$

Replacing $\mathbf{x}$ in previous expression leads to:

$$\begin{aligned}
&\min_{\mathbf{x}} \frac{1}{2}\|\mathbf{y}-\mathbf{x}\|_2^2 + \lambda TV(\mathbf{x}) \\
=&\lambda \max_{\|\mathbf{z}\|_\infty \leq 1} \frac{\lambda}{2}\|\operatorname{div} \mathbf{z}\|_2^2 - \langle \mathbf{y}, \operatorname{div} \mathbf{z} \rangle - \lambda \|\operatorname{div} \mathbf{z}\|_2^2 \\
=&\lambda \max_{\|\mathbf{z}\|_\infty \leq 1} -\frac{\lambda}{2}\|\operatorname{div} \mathbf{z}\|_2^2 - \langle \mathbf{y}, \operatorname{div} \mathbf{z} \rangle \\
=&-\lambda \min_{\|\mathbf{z}\|_\infty \leq 1} \frac{\lambda}{2}\|\operatorname{div} \mathbf{z}\|_2^2 + \langle \mathbf{y}, \operatorname{div} \mathbf{z} \rangle \\
=&-\frac{1}{2} \min_{\|\mathbf{z}\|_\infty \leq 1} \lambda^2 \|\operatorname{div} \mathbf{z}\|_2^2 + 2\lambda \langle \mathbf{y}, \operatorname{div} \mathbf{z} \rangle \\
=&-\frac{1}{2} \min_{\|\mathbf{z}\|_\infty \leq 1} \|\lambda \operatorname{div} \mathbf{z} + \mathbf{y}\|_2^2 - \|\mathbf{y}\|_2^2 \\
=&-\frac{\lambda^2}{2} \min_{\|\mathbf{z}\|_\infty \leq 1} \|\operatorname{div} \mathbf{z} + \frac{\mathbf{y}}{\lambda}\|_2^2 - \frac{1}{\lambda^2}\|\mathbf{y}\|_2^2
\end{aligned} \tag{4.18}$$

Hence $\mathbf{z}^*$ is obtained by:

$$\mathbf{z}^* = \operatorname*{arg\,min}_{\|\mathbf{z}\|_\infty \leq 1} \|\operatorname{div} \mathbf{z} + \frac{\mathbf{y}}{\lambda}\|_2^2$$

This provides the result from Chambolle [30].

This constrained problem can be solved by a projected gradient algorithm. The gradient with respect to $\mathbf{z}$ is given by $-\nabla(\operatorname{div} \mathbf{z} + \mathbf{y}/\lambda)$ which gives the following iterative algorithm to solve the ROF problem:

$$\begin{cases}
\mathbf{x}^n = \mathbf{y} + \lambda \operatorname{div} \mathbf{z}^n \\
\mathbf{z}_i^{n+1} = \dfrac{\mathbf{z}_i^n + (\tau/\lambda)(\nabla \mathbf{x}^n)_i}{\max(1, |\mathbf{z}_i^n + (\tau/\lambda)(\nabla \mathbf{x}^n)_i|)}
\end{cases} \tag{4.19}$$

Where τ is the gradient step. In order to guarantee the algorithm convergence one needs to have:

$$\tau \leq \frac{2}{|||\operatorname{div} \nabla|||}$$

where $|||\operatorname{div} \nabla|||$ stands for the spectral norm of the operator.

Note that Chambolle proposes an alternative strategy based on a fixed point method to

solve this constrained problem, but, our experience is that the fixed-point method does not actually provide faster convergence rates than the simple projected gradient method. Like in any standard gradient descent, the projected gradient just detailed can be improved with a multistep approach [1].

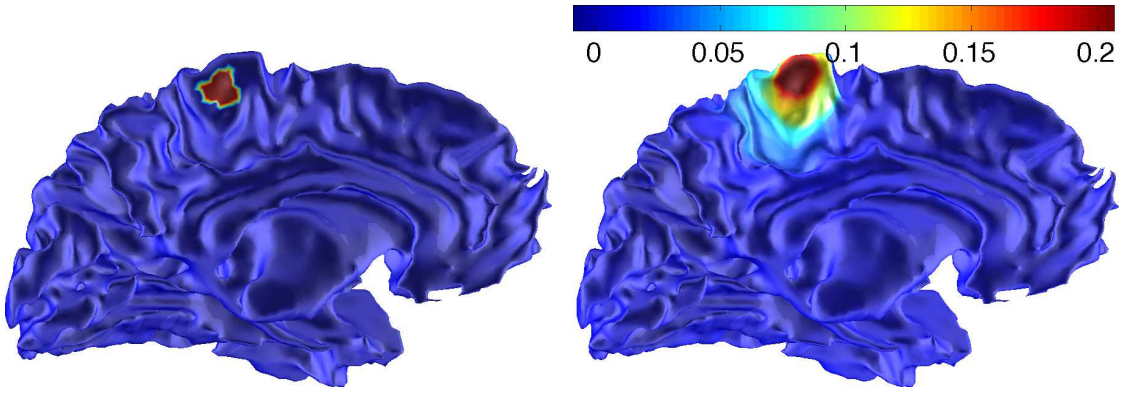
The ROF problem for which we just described an optimization scheme corresponds to the proximity operator associated with the Total Variation penalization. The real solution can then be obtained using forward-backward proximal iterations (algorithm 4.3) or Nesterov iterations (algorithm 4.4).

To our knowledge, Nesterov schemes have never been applied to the inverse problem of M/EEG, in particular with a surface TV prior. In [4], the optimization procedure used suffered from a very slow convergence rate. In this work, Adde *et al*. implemented the forward-backward algorithm 4.3 proposed for TV optimization in [12]. Due to its slow convergence rate, the author suggested to use a fixed step gradient scheme that requires to make the TV prior differentiable by replacing the TV by:

$$TV(\mathbf{x}) = \sum_p \sqrt{(\nabla_p^x \mathbf{x})^2 + (\nabla_p^y \mathbf{x})^2 + (\nabla_p^z \mathbf{x})^2 + \epsilon}, \epsilon > 0 \ .$$

Unfortunately, this scheme does not reach an optimal solution. However, in practice, it provides visually acceptable results in a small amount of time. The Nesterov scheme that we propose here is a principled algorithm to solve (4.13) with guarantees of speed and optimality.

Figure 4.4 presents a result obtained with a TV prior with constrained orientations.



(a) Synthetic active region.

(b) Reconstruction result obtained with a TV prior.

Figure 4.4: Simulation result using a TV prior. (a) The synthetic active region used to simulate MEG measurements. The measurements where corrupted with a small additive Gaussian white noise (SNR=10). The activation pattern was designed to be piecewise constant which is what is adapted for TV minimization. (b) The reconstruction result obtained by solving the inverse problem with a TV prior. The solution presents a clear "hot spot" at the correct location and sets to 0 the majority of the remaining cortical surface.

The TV prior offers a principled to way to cope with spatially extended activations. However good care should be taken when applied with M/EEG data. Our experience shows that when applied with too big values of λ, the TV prior tends to move the active regions towards "flat" cortical regions. The complexe shape of the cortical mantle near the active region may therefore in practice lead to localization errors. Also, to our knowledge, the bias towards superficial sources has never been tackled with a TV prior. And finally, the case of dipolar source spaces with unconstrained orientations has to our knowledge never been treated.

We are now done with the presentation of image-based inverse solvers that work on an

instant-by-instant basis. We will know present solvers that make use of the temporal information in the data.

4.4 SPARSITY AND SPATIOTEMPORAL DATA

The reason for using temporal information in the inverse problem seems relatively natural. Two neighboring time instants carry a very similar information since the underlying physiological phenomena have a low frequency compared to the sampling rate of the recordings. The noise that corrupts the measurements is also highly correlated in time. And our knowledge about physiology favors a vision of the sources as a set of sources, of limited number, whose activity is stable in time.

4.4.1 VESTAL

The VESTAL method [108] addresses the main critic that is made about MCE. When using MCE during a small time window, the set of active dipoles can vary significantly between two neighboring instants, although one would expect a very similar current distribution. The MCE solver leads to "spiky" estimated time courses.

In order to fix this problem, the VESTAL solver proposes to run an ℓ_1 inverse problem at each time instant, like MCE, but projects the sample-wise l1-norm estimates into the signal subspace defined by a set of temporal basis functions. Let us write the SVD of the measurements, $\mathbf{M} = \mathbf{U}\mathbf{S}\mathbf{V}^T$. The first columns of $\mathbf{V}$ are used as temporal basis functions.

The critiques that can be made about the VESTAL solver are the following. First, the optimization scheme proposed in [108] can be largely improved, using the LARS algorithm for example. Second, the authors do not clearly state what cost function is actually optimized by their procedure. The following algorithm proposes a much better approach with the same objective of mixing time and space with a sparsity inducing prior.

4.4.2 ℓ_1 over space and ℓ_2 over time

An ℓ_1 prior brings sparsity, *i.e.,* a limited number of active sources, while an ℓ_2 prior brings what we call diversity, *i.e.,* no zero coefficients. The ℓ_2 norm spreads the energy over all the sources and therefore brings smoothness. By using an ℓ_1 prior over space while keeping an ℓ_2 prior over time, the problem of "spiky" activation times series faced by the MCE method is addressed. We recall that the MCE solver runs on each time instant independently with an ℓ_1 prior.

The approach described above consists in penalizing the inverse problem with a mixed norm given by:

$$\|\mathbf{X}\|_{21} = \sum_i \sqrt{\sum_t x_{it}^2} \ , \tag{4.20}$$

where i indexes space and t indexes time. One will denote this norm ℓ_{21}.

It can also be modified to take into account some weighting coefficients:

$$\|\mathbf{X}\|_{w;21} = \sum_i \sqrt{\sum_t w_i x_{it}^2} \tag{4.21}$$

in order to reduce the bias for superficial sources.

The optimization problem becomes:

$$\mathbf{X}^* = \arg\min_{\mathbf{X}} \frac{1}{2}\|\mathbf{M} - \mathbf{G}\mathbf{X}\|_2^2 + \lambda\|\mathbf{X}\|_{w;21}, \lambda > 0 \ . \tag{4.22}$$

This norm introduces an ℓ_1 norm at a group level. Within a group the coefficients are compared using an ℓ_2 norm. In statistics and machine learning, it is known as the *Group-LASSO* problem[235]. We also refer the reader to [141] for more details on mixed norms in the signal processing context.

In the M/EEG community this norm was recently proposed in this exact form by Ou *et al*. in [166]. Prior to this work, [74] proposed another grouping strategy this time between orientations. In [166], the authors detail how to group both time and orientations. If the coefficients of $\mathbf{X}$ are indexed by the position i, the orientation r, and the time t, the corresponding prior is given:

$$\|\mathbf{X}\|_{21} = \sum_i \sqrt{\sum_r \sum_t x_{irt}^2} \ .$$

Integrating the orientations in the prior does not add any difficulty in the optimization.

In order to speed up the computation, it is proposed in [166] to reduce the rank of the matrix $\mathbf{M}$ with an SVD and to work on the SVD components rather than on the raw temporal data. When considering K SVD components, the size of the matrix to invert is $d_m \times K$ rather than $d_m \times d_t$. Due to the ℓ_2 norm within the groups, this is perfectly justified. If we use all the temporal components of the SVD, we project the data using a basis of orthogonal vectors and the temporal ℓ_2 norm is not changed.

However, in [166], the authors use an interior point methods to solve the optimization problem. Interior point method, also referred to as barrier methods, are a certain class of algorithms to solve linear and nonlinear convex optimization problems. They guarantee the optimality of the solution in polynomial time but do not scale very well to large problems, *i.e.,* source spaces with a high number of dipoles and a lot of time instants. The complexity of the algorithm proposed in [166] is in fact cubic in the number of variables.

As for the ℓ_1 problem, various methods exist to compute the inverse problem with an ℓ_{21} norm. The particular method that does not apply is the LARS. One can use a coordinate descent algorithm [77], an IRLS solver similar to the one proposed for the LASSO or an iterative method based on proximity operators similar to what has been presented for the ℓ_1 and TV priors [55, 224].

To apply the forward-backward iterations or the optimal scheme from Nesterov, one needs to compute the proximity operator associated to a $\|\cdot\|_{21}$ prior.

Let us denote by $\mathbf{X}^i$ the ith row of $\mathbf{X}$. By definition, the proximity operator associated to the ℓ_{21} norm is given:

$$\begin{aligned} \mathbf{X}^* &= \text{prox}_{\lambda\|\ \|_{21}}(\mathbf{Y}) \\ &= \arg\min_{\mathbf{X}} \frac{1}{2}\|\mathbf{Y} - \mathbf{X}\|_F^2 + \lambda\|\mathbf{X}\|_{21} \end{aligned} \tag{4.23}$$

The solution is obtained using the following proposition:

Proposition 4.4 (Group-LASSO proximity operator)**.** *The solution* $\mathbf{X}^* = \text{prox}_{\lambda\|\ \|_{21}}(\mathbf{Y})$ *of the proximity operator associated to the Group-LASSO is given group by group (here row by row) by:*

$$\mathbf{X}^{i*} = \mathbf{Y}^i \left(1 - \frac{\lambda}{\|\mathbf{Y}^i\|_2}\right)^+ \ . \tag{4.24}$$

Using the algorithms 4.4 and 4.3 detailed for the ℓ_1 prior, one can compute the optimization in (4.22). The only difference comes from the modification of the proximity operator. If order to run this algorithm with EMBAL , the code snippet in table 4.6 can be used.

```
clear options
options.penalty = '121'; % Use L21 as Prior
options.project = 5; % Set K=5 (SVD of the measurements)
options.lambda = 1e-5;
X = prox_inverse(M,G,options);
```

Table 4.6: Runing sparse inverse modeling with temporal data using proximity operators and EMBAL .

This strategy that consists in grouping time instants, achieves the same goal as the VESTAL solver. It provides sparse solution over space and smooth solutions over time. However, the methodology behind the Group-LASSO is much more principled and is expected to provide better results.

Due to the use of an ℓ_1 prior over space, that favors focal sources with low SNR, this approach is not adapted to the reconstruction of spatially extended activations. Another critic that can be addressed to this approach is that the active set, *i.e.,* the list of active dipoles, is the same for the full time window of interest. This also implies that the length of the time window considered has an influence on the active set and therefore the solution. The ℓ_1 prior implies that an optimal solution has fewer active sources than the number of sensors. By extending the size of the time window to look for example at late responses, the active set estimated just after stimulation might be changed. Indeed, the ℓ_2 prior over time, and the grouping of all time instants, implies that if a dipole has an activation at one time instant, it has one during the full time period. There is no way with this method to see a dipole become active in the middle of the time window of interest. It is however possible to handle such a case, by using groups with overlaps. In the current setting, a variable x_{it} belongs to only one group, the ith. Setting the ith group to zero sets the activation of dipole i to zero for the full time period. Introducing overlap between groups means allowing a variable x_{it} to belong to multiple groups. This would circumvent the limitations of the ℓ_{21} prior. Unfortunately, the optimization with overlapping groups is not trivial to handle. We refer the reader to [116] and [112] for more details on this topic.

4.5 Sparse priors with multiple experimental conditions: ℓ_{212}

In the previous paragraph, we have seen how the use of priors with mixed norms, like ℓ_{21}, allows to integrate *structured sparsity*. From a Bayesian point of view, using such a mixed norm introduces some coupling between coefficients, instead of the independency hypothesis associated with an ℓ_p norm. By grouping time instants, Ou *et al*. [166] introduced a coupling between all time instants.

We now propose to integrate in a single framework a prior that brings structured sparsity between space, time but also the experimental condition. The development of this method was motivated by its application to the retinotopic mapping with MEG presented in chapter 5.

During an experiment, a subject is generally asked to perform different cognitive tasks or to respond to various external stimuli. They are referred as different experimental conditions. With a standard ℓ_2 prior, it may occur, that the estimated active cortical regions in condition 1 overlap the active regions of condition 2, which may often be unrealistic considering what is known about neuroanatomy. In order to take into account this anatomical knowledge, and obtain more accurate mappings of some brain functional organization, we propose to use a prior that penalizes overlap between active regions.

4.5.1 Method

In order to introduce such inter-condition sparsity constraints, currents corresponding to all conditions have to be estimated simultaneously. Let d_k denote the number of conditions. It is achieved by concatenating all measurements, $\mathbf{M} \in \mathbb{R}^{d_m \times d_k d_t}$. Let $\mathbf{X} \in \mathbb{R}^{d_x \times d_k d_t}$ have its elements now indexed by (i,k,t), i indexes space, k the condition and t the time.

In chapter 5, the following norm is used with Fourier coefficients. Therefore, the following definition and the following proposition, that gives the associated proximity operator, is written with complex valued coefficients.

Definition 4.3 (Three level mixed norm)**.** Let $\mathbf{x} \in \mathbb{C}^{d_x d_k d_t}$ be indexed by a triple index (i,k,t) such that $\mathbf{x} = (x_{i,k,t})$. Let $p,q,r \geq 1$ and $\mathbf{w} \in \mathbb{R}_{+,*}^{d_x d_k d_t}$ be a sequence of strictly positive weights labelled by a triple index (i,k,t). We call mixed norm of $\mathbf{x}$ the norm $\ell_{\mathbf{w};p,q,r}$ defined by

$$\|\mathbf{x}\|_{\mathbf{w};pqr} = \left(\sum_{i=1}^{d_x} \left(\sum_{k=1}^{d_k} \left(\sum_{t=1}^{d_t} w_{i,k,t} |x_{i,k,t}|^p \right)^{q/p} \right)^{r/q} \right)^{1/r} .$$

The problem that is addressed here is:

$$\mathbf{X}^* = \arg\min_X \|\mathbf{M} - \mathbf{G}\mathbf{X}\|_F^2 + \lambda \|\mathbf{X}\|_{\mathbf{w};212}^2 \, , \lambda \in \mathbb{R}_+ \ . \tag{4.25}$$

A ℓ_1 prior is set over the index k corresponding to the condition, while an ℓ_2 prior is used over space and time. By doing so, each dipole has an incentive to explain a small number of conditions. The conditions are not supposed to change during the time window. Note that $\|\mathbf{X}\|_{\mathbf{w};222} = \|\mathbf{X}\|_{\mathbf{w};F}$ and that if $d_k = 1$, *i.e.,* only one condition, $\|\mathbf{X}\|_{\mathbf{w};212} = \|\mathbf{X}\|_{\mathbf{w};F}$. This means that in the case where only one condition is considered, the $\ell_{\mathbf{w};212}$ solution corresponds to the widely used MN and WMN (see section 3.2).

Here also, solving (4.25) is based on the computation of the proximity operator associated to the $\ell_{\mathbf{w};212}$ norm.

The proximity operator associated with the mixed norm $\|.\|_{\mathbf{w};212}^2$ is analytically given by the following proposition. We denote $\mathbf{y}_{i,k,\bullet} = (y_{i,k,1}, y_{i,k,2}, \ldots, y_{i,k,T})$.

Proposition 4.5. *Let $\mathbf{y} \in \mathbb{C}^{d_x d_k d_t}$ be indexed by a triple index (i,k,t). Let $\mathbf{w}$ a sequence of strictly positive weights such that $\forall t,\ w_{i,k,t} = w_{i,k}$. Let $r_{i,k}$ be defined as $r_{i,k} \overset{def}{=} \|\mathbf{y}_{i,k,\bullet}\|_{\mathbf{w},2} / w_{i,k}$, where $\|\mathbf{y}_{i,k,\bullet}\|_{\mathbf{w},2} \overset{def}{=} \sqrt{w_{i,k} \sum_t |y_{i,k,t}|^2}$. For each i, let the indexing denoted by k_i' be defined such that $\forall k_i', r_{i,k_i'+1} \leq r_{i,k_i'}$. Let the index K_i and the quantity $K_{\mathbf{w}_i} \overset{def}{=} \sum_{k_i=1}^{K_i} w_{i,k_i}$ be defined such that*

$$\lambda \sum_{k_i'=1}^{K_i} w_{i,k_i'} \left(r_{i,k_i'} - r_{i,K_i} \right) < r_{i,K_i} \leq \lambda \sum_{k_i'=1}^{K_i+1} w_{i,k_i'} \left(r_{i,k_i'} - r_{i,K_i} \right) \ .$$

Then the solution $\mathbf{z} = \mathrm{prox}_{\frac{\lambda}{2}\|.\|_{\mathbf{w};212}^2}(\mathbf{y})$ is given for each coordinate (i,k,t) by

$$z_{i,k,t} = y_{i,k,t} \left(1 - \frac{\lambda \sqrt{w_{i,k}}}{1 + \lambda K_{\mathbf{w}_i}} \frac{\sum_{k_i'=1}^{K_i} \|\mathbf{y}_{i,k_i',\bullet}\|_{\mathbf{w},2}}{\|\mathbf{y}_{i,k,\bullet}\|_2} \right)^+ .$$

Proof. To simplify the notations in the demonstration, we remove the $\frac{1}{2}$ in the proximity

operator. We address here the equivalent problem:

$$\mathbf{x}^* = \arg\min_{\mathbf{x}} \|\mathbf{y} - \mathbf{x}\|_2^2 + \lambda \|\mathbf{x}\|_{\mathbf{w};212}^2 \ , \tag{4.26}$$

with

$$\|\mathbf{x}\|_{\mathbf{w};212}^2 = \sum_i \left(\sum_j \left(\sum_k w_{i,j} |x_{i,j,k}|^2 \right)^{1/2} \right)^2 \ .$$

To simplify the notations we write $\|\mathbf{y}_{i,k,\bullet}\|_2$ and $\|\mathbf{y}_{i,k,\bullet}\|_{\mathbf{w},2}$ respectively as $\|\mathbf{y}_{i,k}\|_2$ and $\|\mathbf{y}_{i,k}\|_{\mathbf{w},2}$.

Let us derive the functional in (4.26) with respect to $x_{i,j,k}$. It leads to the following system of variational equations:

$$\begin{cases} |x_{i,j,k}| = |y_{i,j,k}| - \lambda \sqrt{w_{i,j}} |x_{i,j,k}| \|\mathbf{x}_{i,j}\|_2^{-1} \|\mathbf{x}_i\|_{\mathbf{w};21} \\ \arg(x_{i,j,k}) = \arg(y_{i,j,k}) \end{cases}$$

which gives:

$$\begin{aligned} |x_{i,j,k}| \left(1 + \lambda \sqrt{w_{i,j}} \|\mathbf{x}_{i,j}\|_2^{-1} \|\mathbf{x}_i\|_{\mathbf{w};21}\right) &= |y_{i,j,k}| \\ \Rightarrow |x_{i,j,k}|^2 \left(1 + \lambda \sqrt{w_{i,j}} \|\mathbf{x}_{i,j}\|_2^{-1} \|\mathbf{x}_i\|_{\mathbf{w};21}\right)^2 &= |y_{i,j,k}|^2 \end{aligned} \tag{4.27}$$

By summing over k, we get:

$$\begin{aligned} \|\mathbf{x}_{i,j}\|_2 \left(1 + \lambda \sqrt{w_{i,j}} \|\mathbf{x}_{i,j}\|_2^{-1} \|\mathbf{x}_i\|_{\mathbf{w};21}\right) &= \|\mathbf{y}_{i,j}\|_2 \\ \Rightarrow \|\mathbf{x}_{i,j}\|_2 + \lambda \sqrt{w_{i,j}} \|\mathbf{x}_i\|_{\mathbf{w};21} &= \|\mathbf{y}_{i,j}\|_2 \ . \end{aligned} \tag{4.28}$$

We have that:

$$\|\mathbf{x}_i\|_{\mathbf{w};21} = \sum_{l / \|\mathbf{x}_{i,l}\|_2 > 0} \sqrt{w_{i,l}} \|\mathbf{x}_{i,l}\|_2 \ .$$

which implies that:

$$\|\mathbf{x}_{i,j}\|_2 = \|\mathbf{y}_{i,j}\|_2 - \lambda \sqrt{w_{i,j}} \sum_{k / \|\mathbf{x}_{i,k}\|_2 > 0} \sqrt{w_{i,k}} \|\mathbf{x}_{i,k}\|_2 \ . \tag{4.29}$$

Using (4.28), we have, if j and k satisfy $\|\mathbf{x}_{i,j}\|_2 > 0$ and $\|\mathbf{x}_{i,k}\|_2 > 0$, that:

$$\frac{\|\mathbf{x}_{i,k}\|_2}{\sqrt{w_{i,k}}} = \frac{\|\mathbf{x}_{i,j}\|_2}{\sqrt{w_{i,j}}} + \frac{\|\mathbf{y}_{i,k}\|_2}{\sqrt{w_{i,k}}} - \frac{\|\mathbf{y}_{i,j}\|_2}{\sqrt{w_{i,j}}} \ .$$

By injecting it in (4.29) we get:

$$\begin{aligned} \|\mathbf{x}_{i,j}\|_2 &= \|\mathbf{y}_{i,j}\|_2 - \lambda \sqrt{w_{i,j}} \sum_{k / \|\mathbf{x}_{i,k}\|_2 > 0} w_{i,k} \left(\frac{\|\mathbf{x}_{ij}\|_2}{\sqrt{w_{i,j}}} + \frac{\|\mathbf{y}_{i,k}\|_2}{\sqrt{w_{i,k}}} - \frac{\|\mathbf{y}_{i,j}\|_2}{\sqrt{w_{i,j}}} \right) \\ \Leftrightarrow \|\mathbf{x}_{i,j}\|_2 &= \|\mathbf{y}_{i,j}\|_2 - \lambda K_{\mathbf{w}_i} \|\mathbf{x}_{i,j}\|_2 + \lambda K_{\mathbf{w}_i} \|\mathbf{y}_{i,j}\|_2 - \lambda \sqrt{w_{i,j}} \sum_{k / \|\mathbf{x}_{i,k}\|_2 > 0} \sqrt{w_{i,k}} \|\mathbf{y}_{i,k}\|_2 \ , \\ \Leftrightarrow \|\mathbf{x}_{i,j}\|_2 &= \|\mathbf{y}_{i,j}\|_2 - \frac{\lambda \sqrt{w_{i,j}}}{1 + \lambda K_{\mathbf{w}_i}} \sum_{k / \|\mathbf{x}_{i,k}\|_2 > 0} \|\mathbf{y}_{i,k}\|_{\mathbf{w};2} \end{aligned}$$

where $K_{\mathbf{w}_i} = \sum_{k / \|\mathbf{x}_{i,k}\|_2 > 0} w_{i,k}$. This provides the solution for $\|\mathbf{x}_{i,j}\|_2 > 0$. When

$$\|\mathbf{y}_{i,j}\|_2 - \frac{\lambda \sqrt{w_{i,j}}}{1 + \lambda K_{\mathbf{w}_i}} \sum_{k / \|\mathbf{x}_{i,k}\|_2 > 0} \|\mathbf{y}_{i,k}\|_{\mathbf{w};2} \leq 0$$

it implies that $\|\mathbf{x}_{i,j}\|_2 = 0$.

We therefore have for all i:

$$\|\mathbf{x}_{i,j}\|_2 = \left(\|\mathbf{y}_{i,j}\|_2 - \frac{\lambda\sqrt{w_{i,j}}}{1+\lambda K_{\mathbf{w}_i}} \sum_{k/\|\mathbf{x}_{i,k}\|_2>0} \|\mathbf{y}_{i,k}\|_{\mathbf{w};2} \right)^+ .$$

Let us introduce the indexing k'_i such that $\|\mathbf{x}_{i,k'_i+1}\|_2 \leq \|\mathbf{x}_{i,k'_i}\|_2$ and the index K_i such that $\|\mathbf{x}_{i,K_i}\|_2 > 0$ and $\|\mathbf{x}_{i,K_i+1}\|_2 = 0$. After reordering, we have that:

$$\|\mathbf{y}_{i,K_i}\|_2 - \frac{\lambda\sqrt{w_{i,K_i}}}{1+\lambda K_{\mathbf{w}_i}} \sum_{k'_i=1}^{K_i} \|\mathbf{y}_{i,k'_i}\|_{\mathbf{w};2} > 0 \ ,$$

which leads to:

$$\lambda \sum_{k'_i=1}^{K_i} w_{i,k'_i} \left(r_{i,k'_i} - r_{i,K_i} \right) < r_{i,K_i} \ .$$

This provides the reordering in the proposition and the equation:

$$\|\mathbf{x}_{i,j}\|_2 = \left(\|\mathbf{y}_{i,j}\|_2 - \frac{\lambda\sqrt{w_{i,j}}}{1+\lambda K_{\mathbf{w}_i}} \sum_{k'_i=1}^{K_i} \|\mathbf{y}_{i,k'_i}\|_{\mathbf{w},2} \right)^+ . \tag{4.30}$$

Let us rewrite (4.27):

$$\begin{aligned} |x_{i,j,k}| &= \frac{|y_{i,j,k}|}{1+\lambda\sqrt{w_{i,j}}\|\mathbf{x}_{i,j}\|_2^{-1}\|\mathbf{x}_i\|_{\mathbf{w};21}} \\ &= \frac{|y_{i,j,k}|\|\mathbf{x}_{i,j}\|_2}{\|\mathbf{x}_{i,j}\|_2 + \lambda\sqrt{w_{i,j}}\|\mathbf{x}_i\|_{\mathbf{w};21}} \end{aligned}$$

Using (4.28), we get:

$$|x_{i,j,k}| = \frac{|y_{i,j,k}|\|\mathbf{x}_{i,j}\|_2}{\|\mathbf{y}_{i,j}\|_2} \ .$$

By injecting the result in (4.30) in this equation we get:

$$\begin{aligned} |x^*_{i,j,k}| &= \frac{|y_{i,j,k}| \left(\|\mathbf{y}_{i,j}\|_2 - \frac{\lambda\sqrt{w_{i,j}}}{1+\lambda K_{\mathbf{w}_i}} \sum_{j=1}^{K_i} \|\mathbf{y}_{i,j}\|_{\mathbf{w},2} \right)^+}{\|\mathbf{y}_{i,j}\|_2} \\ &= |y_{i,j,k}| \left(1 - \frac{\lambda\sqrt{w_{i,j}}}{1+\lambda K_{\mathbf{w}_i}} \frac{\sum_{j=1}^{K_i} \|\mathbf{y}_{i,j}\|_{\mathbf{w},2}}{\|\mathbf{y}_{i,j}\|_2} \right)^+ . \end{aligned}$$

□

Remarks.

1. If $T = 1$, then $\text{prox}_{\lambda\|.\|^2_{\mathbf{w};212}}(\mathbf{y}) = \text{prox}_{\lambda\|.\|^2_{\sqrt{\mathbf{w}};12}}(\mathbf{y})$ which corresponds to the Elitist-Lasso problem [129].

2. This proposition also provides the proximity operator for the $\ell_{\mathbf{w};21}$ norm introduced in section 4.4.2.

3. The proximity operator is known analytically. It is simply a shrinkage operator after a sorting operation. It implies that the solution is exact and relatively fast to compute.

Columns $(G_{\cdot i})_i$ of M/EEG forward operators are not normalized. The closer is the dipole i from the head surface, the bigger is $\|G_{\cdot i}\|_2$. This implies that a naive inverse procedure would favor dipoles close to the head surface. Using a weighted norm is an alternative to cope with this problem. With the mixed norm $\|.\|_{\mathbf{w},212}$, it is done by setting $w_{i,k} = w_i = \|G_{\cdot i}\|_2$.

$\ell_{\mathbf{w};212}$ with unconstrained orientations

Up to here, the $\ell_{\mathbf{w};212}$ prior was presented assuming a source space with constrained orientations. It is however easy to integrate into the framework the unconstrained case. Let us denote d_r the number of oriented dipoles set at each brain location. The number of brain locations is denoted d_x. Let $\mathbf{X} \in \mathbb{R}^{d_x d_r \times d_k d_t}$ have its elements now indexed by (i, r, k, t), i indexes the position, r the orientation, k the condition and t the time. The prior is, in this case, given by:

$$\|\mathbf{X}\|_{\mathbf{w};212} = \left(\sum_{i=1}^{d_x} \left(\sum_{k=1}^{d_k} \sqrt{\sum_{t=1}^{d_t} \sum_{r=1}^{d_r} w_i |x_{i,r,k,t}|^2} \right)^2 \right)^{1/2} .$$

This norm is still of the form of $\ell_{\mathbf{w};212}$ and leads therefore to the same optimization strategy.

Like in the constrained case, bias for superficial sources can be corrected by setting:

$$w_i = \sqrt{\sum_r \|G_{\cdot ir}\|_2^2} ,$$

where $G_{\cdot ir}$ stands for the forward field of the dipole with position i and orientation r.

4.5.2 Simulations

By setting a ℓ_1 prior between conditions, the mixed norm proposed penalizes overlap between active cortical regions. In order to illustrate this, we generated two synthetic datasets. The first reproduces part of the organization of the primary somatosensory cortex (S1) [174]. Three, non overlapping, cortical regions with a similar area (cf. Fig. 4.6a), that could correspond to the localization of 3 right hand fingers, have been computed and used to generate synthetic measurements corrupted with an additive Gaussian random noise. The amplitude of activation for the most temporal region (colored in red in Fig. 4.6), that could correspond to the thumb, was set twice bigger than the amplitudes of the two other regions. This situation, where the source amplitudes differ between conditions, is relatively common with real M/EEG data. The inverse problem was then computed with a standard $\|.\|_{\mathbf{w};F}$ norm and the $\|.\|_{\mathbf{w};212}$ mixed norm. Within the 3 neighboring active regions, a label corresponding to the condition giving the maximum of amplitude in each of the three conditions was assigned to each dipole. Quantification of performance was done for multiple values of signal-to-noise ratio (SNR) by counting the percentage of dipoles that have been incorrectly labeled. The SNR is defined here as 20 times the log of the ratio between the norm of the signal and the norm of the added noise. Results are also presented in Fig. 4.5. Results with an ℓ_1 prior, has also been added. It can be observed that the $\|.\|_{\mathbf{w};212}$ produces systematically the best result. The ℓ_1 is very rapidly affected by the decrease of SNR, which is known in the M/EEG community. In order to have a fair comparison between all methods, λ was set in each case to have $\|\mathbf{M} - \mathbf{G}\mathbf{X}^*\|_F$ equal to the norm of the added noise, known in the simulations.

Results are illustrated in Fig. 4.6b and 4.6c on a region of interest (ROI) around the left primary somatosensory cortex. It can be observed that the extent of the most lateral region, obtained with $\|.\|_{\mathbf{w};F}$, is overestimated while the result obtained with the $\|.\|_{\mathbf{w};212}$ mixed norm is relatively accurate. Similar simulations have been performed in the primary visual cortex (V1), reproducing the well known retinotopic organization of V1. Results are presented in

Fig. 4.7. Simulations lead to the same conclusion about the superiority of the $\|.\|_{\mathbf{w};212}$ mixed norm for the mapping of such brain functional organizations.

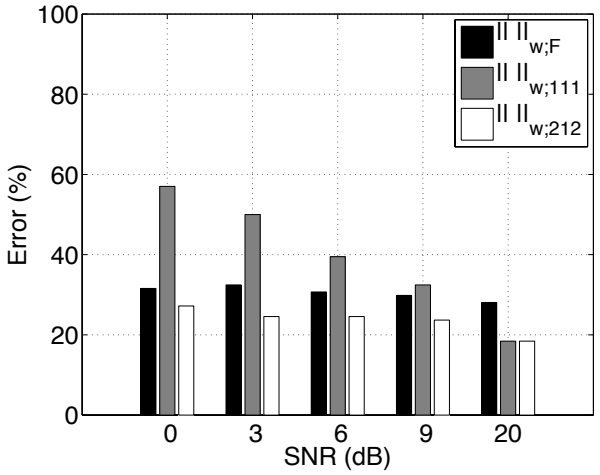
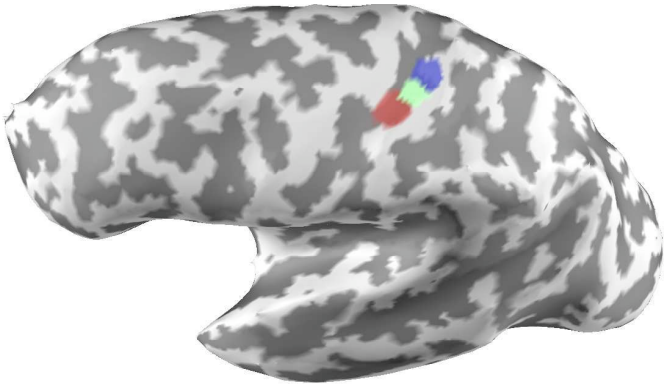
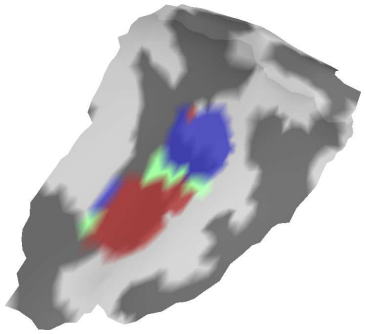


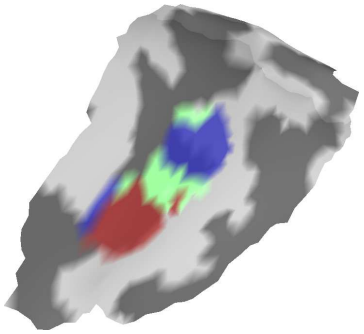
Figure 4.5: Evaluation of $\|.\|_{\mathbf{w};F}$ vs. $\|.\|_{\mathbf{w};212}$ vs. $\|.\|_{\mathbf{w};111}$ estimates on synthetic somatosensory data. The error represents the percentage of wrongly labeled dipoles.



(a) Simulation data

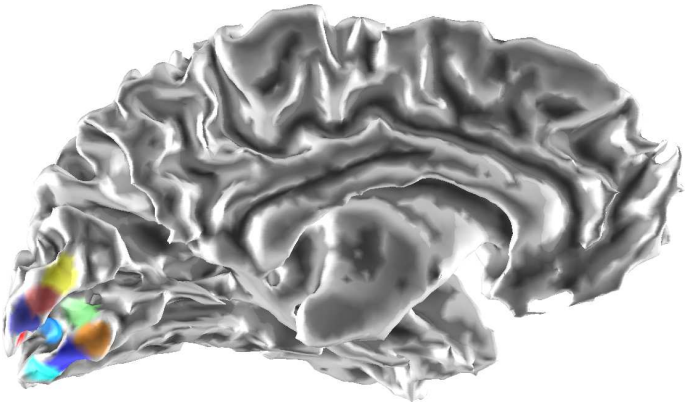


(b) $\|.\|_{\mathbf{w};F}$ result (ROI)

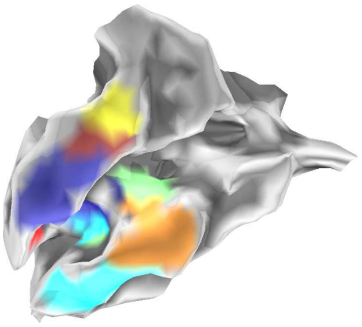


(c) $\|.\|_{\mathbf{w};212}$ result (ROI)

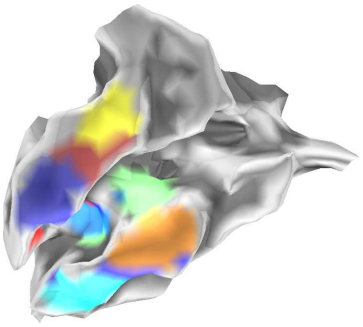
Figure 4.6: Illustration of result on the primary somatosensory cortex (S1) (SNR = 20dB). Neighboring active regions reproduce the organization of S1.



(a) Simulation data



(b) $\|.\|_{\mathbf{w};F}$ result (ROI)



(c) $\|.\|_{\mathbf{w};212}$ result (ROI)

Figure 4.7: Illustration of result on the primary visual cortex (V1) with $\mathrm{SNR} = 20\mathrm{dB}$. Neighboring active regions reproduce the retinotopic organization of V1. When comparing the two results in (b) and (c) with the simulation data in (a), it can be observed that the result in (c) obtained with the $\|.\|_{\mathbf{w};212}$ prior provides the most accurate result.

4.5.3 MEG study

Results of the proposed algorithm using MEG data from a somatosensory experiment are now presented. The data acquisition was done using a CTF Systems Inc. Omega 151 system with a 1250 Hz sampling rate. The somatosensory stimulation was an electrical square-wave pulse delivered randomly to the thumb, index, middle and little finger of each hand of a healthy right-handed subject. Evoked data were computed by averaging 400 repetitions of the stimulation of each finger. To produce precise localization results, the triangulation over which cortical activations have been estimated was sampled with a very high number of vertices (about 55 000). The forward modeling was performed with a spherical head model[1] using dipoles with fixed orientations given by the normals to the cortex [50].

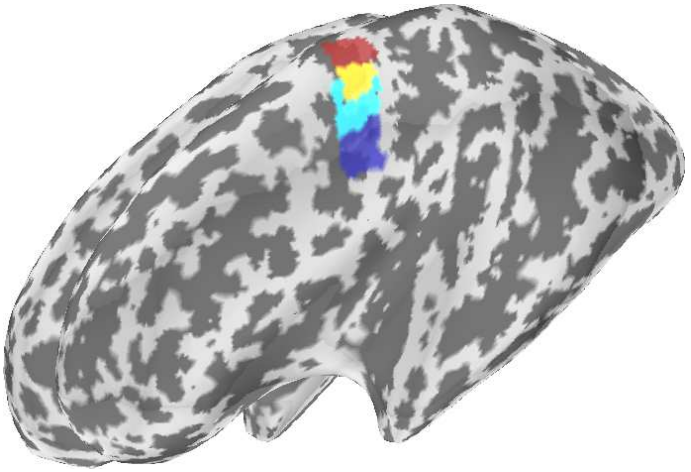
Prior to the current estimation, data were whitened using the noise covariance matrix $\mathbf{\Sigma}$, estimated on the period before stimulation. Let $\mathbf{\Sigma} = \mathbf{L}^T\mathbf{L}$ the Cholesky factorization of $\mathbf{\Sigma}$. Whitening consists in replacing $\mathbf{G}$ by $\mathbf{L}^{-1}\mathbf{G}$ and $\mathbf{M}$ by $\mathbf{L}^{-1}\mathbf{M}$. With an additive Gaussian noise model this implies that the noise, given by $\mathbf{M} - \mathbf{G}\mathbf{X} \in \mathbb{R}^{d_x \times d_k d_t}$, is assumed to have a standard normal distribution. This implies that a good estimate of $\|\mathbf{M} - \mathbf{G}\mathbf{X}^*\|_F$ is given by $\sqrt{d_x d_k d_t}$. Therefore, the regularization parameter λ was set in order for $\mathbf{X}^*$ to be also the solution of the constrained problem: $\mathbf{X}^* = \arg\min_{\mathbf{X}} \|\mathbf{X}\|$ subject to $\|\mathbf{M} - \mathbf{G}\mathbf{X}\|_F \leq \sqrt{d_x d_k d_t}$. The optimization is done with the algorithm 4.5.

Results obtained with the right hand fingers during the period between 42 and 46 ms are presented in figure 4.8. Knowing that for this somatosensory dataset active parcels should

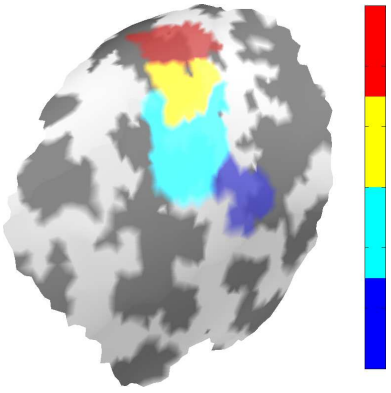
[1] http://neuroimage.usc.edu/brainstorm/

have negative activations around 45 ms, regions with positive activations were first removed. Within the remaining regions, a label was assigned to each dipole based on its maximum amplitude across conditions. For each condition, equivalently each label, the biggest connected component was kept. Each of the 4 estimated components, corresponding to the 4 right hand fingers are presented in Fig. 4.8. Solutions using both norms $\|.\|_{\mathbf{w};F}$ and $\|.\|_{\mathbf{w};212}$ are detailed. With $\|.\|_{\mathbf{w};212}$ the well known organization of the primary somatosensory cortex [174] is successfully recovered, while with $\|.\|_{\mathbf{w};F}$, the component corresponding to the index finger is overestimated leading to an incorrect localization of the area corresponding to the thumb.

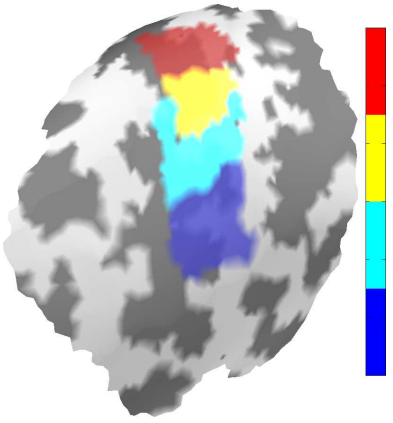
(a) Fingers color coded



(a) $\|.\|_{\mathbf{w},212}$ result



(c) $\|.\|_{\mathbf{w},F}$ result (ROI)



(b) $\|.\|_{\mathbf{w},212}$ result (ROI)

Figure 4.8: Labeling results of the left primary somatosensory cortex in MEG.

4.6 CONCLUSION

In this chapter we presented various approaches and algorithmic details necessary to find the solution of the M/EEG inverse problem with a sparsity inducing prior. State-of-the-art optimization techniques are presented and discussed in order to provide methods tractable on big datasets. We have presented simple algorithms, which is particularly important to facilitate their adoption by the M/EEG community. And finally, we detailed fast algorithms which are of major interest since real studies require to compute statistics in order to validate neuroscientific hypotheses. Indeed, when using robust non-parametric statistics, the inverse solver needs to be run thousands of times. This makes the speed of convergence of the solver a very critical issue.

Our last contribution describes an inter-condition prior that improves the localization of cortical activations by offering the possibility to use a prior between different experimental conditions. By proposing to perform the inverse problem on multiple conditions simultaneously and to use a mixed norm that sets an ℓ_1 prior between each condition, the method penalizes current estimates with an overlap between the corresponding active regions. When such an hypothesis holds anatomically, the more conditions are recorded and used in the inverse problem, the better is the localization of neuronal activity. By keeping an ℓ_2 prior over space and time, the proposed method guarantees a good robustness to noise, like standard ℓ_2 based methods. This approach also improves over the smeared reconstructions observed with standard ℓ_2 inverse solutions. This is confirmed by the simulations and the MEG somatosensory data, with which the method is successfully illustrated.

As for the previous chapter, all the algorithms detailed in this chapter have been implemented and tested with synthetic and real MEG data. The source code of the solvers and the demo scripts with synthetic and real data are available is a Matlab toolbox called EMBAL (Electro-Magnetic Brain Activity Localization):

`https://gforge.inria.fr/projects/embal`

We refer the reader to the demo scripts running on synthetic MEG data:

- demo_inverse_l1.m: contains a comparison of LARS, IRLS, Landweber and Nesterov for the ℓ_1 problem.
- demo_inverse_l21.m: contains a comparison of IRLS, Landweber and Nesterov for the $\ell_{\mathbf{w},21}$ problem.
- demo_inverse_TV.m: contains a comparison of IRLS, simple gradient descent (with adaptive step size obtained with line search), Landweber and Nesterov for the TV problem.

Nothing was said in this chapter about the use of IRLS and simple gradient descent for solving the TV problem as it only solves a *"smoothed"* version of the TV problem. Also, for the IRLS solver, the presence of the gradient makes the solver even more numerically unstable than in the ℓ_p case.

We finish the chapter by listing factors that could be investigated in future studies. Among these is the ability to use a TV prior with dipoles having unconstrained orientations. This can be related to an image deconvolution problem when considering a colored image with 3 channels (red, green and blue). Furthermore, the algorithms detailed in this chapter can be directly applied to multichannel deconvolution problems. Also, we plan to investigate recently proposed iterative algorithms with the same convergence rate as Nesterov's scheme [1] but with a smaller computational complexity. Finally, another major topic we would like to explore is the case where the prior contains overlapping groups of variables. This would allow to deal with temporal data where the active sources might not be flagged as active during the

full time interval. This latter limitation will be addressed in chapter 6 with a completely different approach based on a graph cuts optimization technique.

CHAPTER 5

FAST RETINOTOPIC MAPPING WITH MEG

As demonstrated by research on computer vision in order to reproduce the capabilities of human perception with computers, the human visual system is an amazingly complex machinery. Understanding how visual information is encoded and treated by our brain is still a big challenge that brain functional imaging has been trying to tackle in the last decades thanks to advanced techniques like fMRI [24, 62, 64, 204, 222, 223], PET [75], surface electrodes [234], optical tomography [236], near-infrared spectroscopy [95], EEG [110, 213], and MEG [69, 82, 151, 197].

The motivation for the work presented in this chapter is twofold. First we wanted to investigate how well MEG could reproduce the retinotopic maps obtained by standard protocols in fMRI. Second, we wanted to exploit the excellent temporal resolution of MEG to get an access to brain dynamics during visual processing.

In this chapter we start by presenting the basics of the human visual system and insist particularly on the retinotopic properties of the primary visual cortex (V1). Results obtained in the literature by functional imaging are then presented. We will then describe the experimental protocol designed in order to explore the retinotopic mapping of V1 with MEG. The methods used for data processing, functional mapping and timing include signal extraction with spectral analysis, inverse modelling and statistics with resampling techniques using permutations. Some information on our experimental protocol, analysis methods and results are presented in [2, 42, 44, 45, 46].

Contents

5.1 From the eyes to the cortex

We start this section with a description of the path conveying the visual information from the eyes to the primary visual cortex (V1). We then present briefly the organization of the visual cortex. For more comprehensive description, see for instance [29, 124, 167].

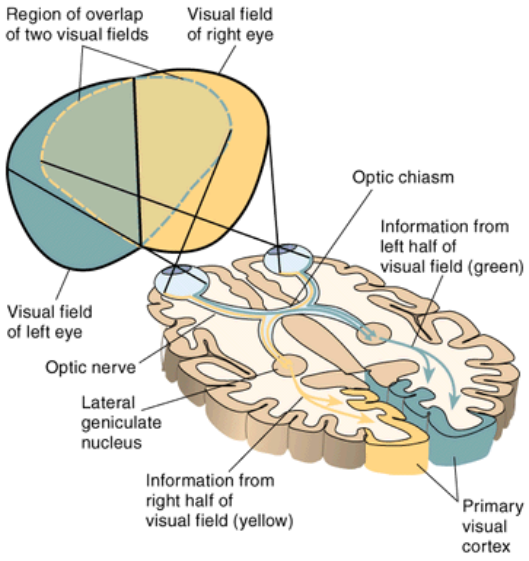


Figure 5.1: The path of the visual information from the eyes to the primary visual cortex (Adapted from `http://homepage.psy.utexas.edu/homepage/Class/Psy308/Salinas/Vision/Vision.html`).

In primates, in particular in humans, the visual system includes many anatomical elements, from the eyes to the cortex. In the eye, light goes successively through the cornea, the aqueous humor, and the pupil. Next it passes through the lens, before entering the vitreous humor. It finally reaches the retina, which is covered with over 125 million photosensitive receptors of two families. The cones form a population of around 8 millions cells. Mainly concentrated in the center of the retina, also known as the fovea, the cones are responsible for chromatic and normal lighting (photopic) condition vision (or photopic). About 120 million rods are found everywhere except in the fovea. They deal with black and white perception and low-lighting conditions (scotopic).

These photosensitive receptors translate lighting information into electrical information, transmitted to the optical nerves via the ganglion cells. The two optical nerves meet, forming the optic chiasm, after which information is transmitted separately for each visual hemifield (separated vertically with respect to the head position): the information from the left (respectively right) parts of both retina and corresponding to the right (left) visual field is brought together to form the left (right) optical tractus (cf. figure 5.1).

The vast majority of the optical tracts fibers get projected to a part of the thalamic sensory relay system, the Lateral Geniculate Nucleus (LGN). Visual signals from the two eyes remain segregated in the LGN which approximately counts 1 million cells corresponding to the number of optical fibers. Finally, the LGN axons form the optic radiations which reach the primary visual cortex (V1), centered around the calcarine fissure (cf. figure 5.2).

V1, also known as Brodmann area 17 (cf. figure 1.9) or “striate cortex” due to its cytoarchitectonic properties, is viewed as the entry of the visual cortex. It receives most outputs of the LGN.

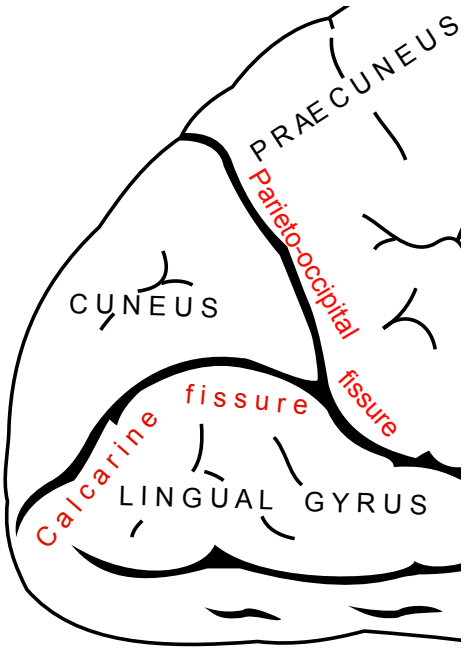


Figure 5.2: Schematic representation of the calcarine fissure in medial view (From 20th U.S. edition of Gray's Anatomy of the Human Body, 1918 (public domain)).

V1 contains a complete (mirror) representation of the contralateral hemifield. This is illustrated by figure 5.3 and figure 5.4. This property corresponds to the *retinotopic* organisation of V1. A practical consequence of this organization, is that besides on patients who underwent damage to the occipital lobe, the left (resp. right) visual hemifield projects to the right (resp. left) occipital cortex.

Beyond this retinotopy, neurons in V1 are organized into sub-regions, each specialized in the analysis of a given visual feature. Among these features are sensitivity to color, contrast, orientations or direction of motion. The sensitivity and selectivity to orientation was first measured by Hubel and Wiesel (Nobel Prize in Physiology or Medicine in 1981) in 1959. By inserting a microelectrode into the primary visual cortex of an anesthetized cat they discovered that some neurons responded more strongly to one particular orientation while other neighboring neurons were more sensitive to other orientations. They called such neurons of V1 "simple cells". This observation motivated the design of the visual stimuli used in the MEG experiment detailed below in section 5.4.

Next to V1 are other visual areas reported in the literature. Without going into much detail, studies on macaque monkeys lead Felleman and Van Essen [70] to differentiate 30 areas based on four main criteria: (i) local cortical cells architecture, (ii) connectivity patterns across areas, (iii) global functional selectivity and (iv) retinotopy. In humans, the last two criteria were successfully used to unveil several areas. We make below a short list of areas neighboring V1.

V2, also called prestriate area, is subdivided in each hemisphere into two parts: V2v (for ventral) and V2d (for dorsal). They respectively represent the upper and lower contralateral quarterfields. Together the four regions provide a complete map of the visual field. Area V2 mainly receives its inputs from V1. Functionally, V2 has many properties in common with V1. Cells are tuned to simple properties such as orientation, spatial frequency and color.

V3, refers to the region of the cortex located immediately next to V2. Like V2, is sub-

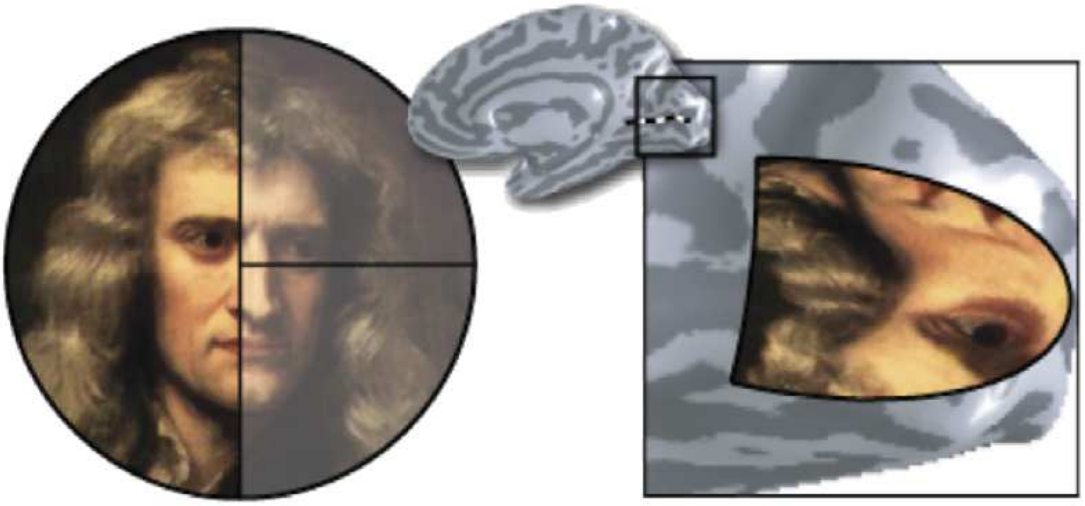


Figure 5.3: Illustration of the retinotopic organization in V1. V1 contains a complete (mirror) representation of the contralateral hemifield (Adapted from [219]).

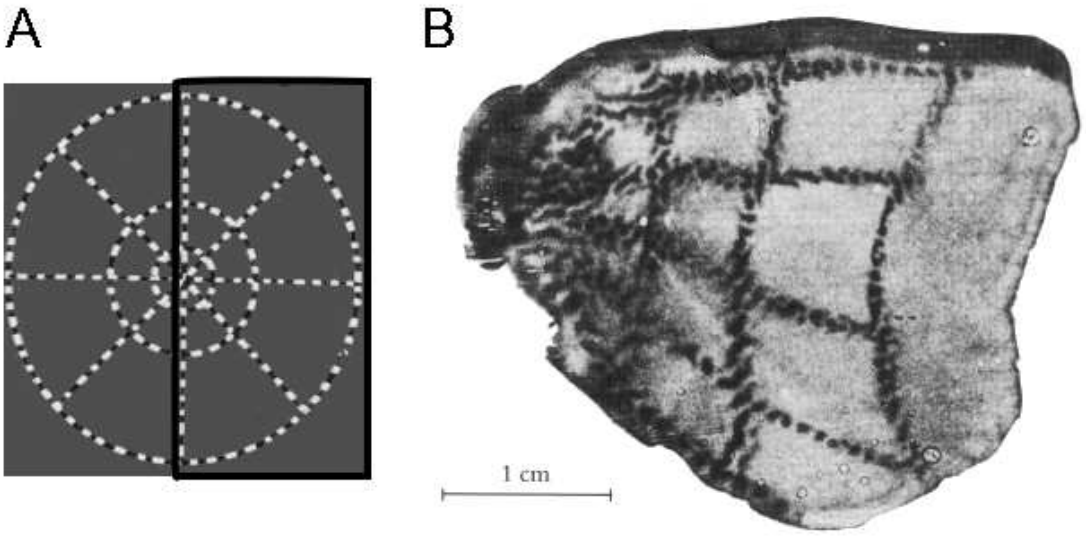


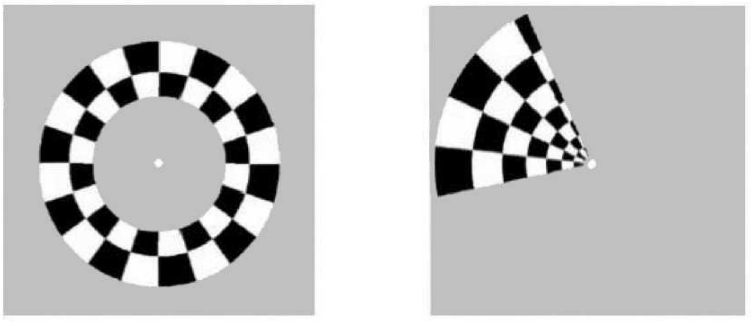
Figure 5.4: Retinotopic organization of the primary visual cortex (V1): The visual field is continuously mapped onto the visual cortex. A: A grid pattern made of con- centric circles is presented to a macaque. The black box corresponds to the right visual field, that will be transferred to the left hemisphere of the brain. B: The grid pattern is reproduced at the level of the primary visual cortex (V1). The mapping from visual field to V1 is termed retinotopic, because it continuously preserves the topology of the retina. The central zone of the visual field is processed with great precision: It is processed by many neurons in V1, comparatively to the periphery. Mathematically, the mapping from visual field to V1 is well approximated by a log- polar scheme (Adapted from [205]).

divided into two parts: V3v ventrally (sometimes also called VP, in reference to the Ventral Posterior area in monkeys) and V3d dorsally. However, contrarily to V1 and V2, there is still some controversy regarding its exact extent and its functional selectivity. Even if no consensus exists for humans, we can consider that V3v represents the upper quadrant and V3d represents the lower quadrant.

5.2 Retinotopic mapping with fMRI

By looking at figure 5.4, it can be noticed that a natural coordinate system for the spatial organization of V1 is the polar coordinate system. This observation led to the design of the stimuli called *wedges* and *rings*. The wedges encode the polarity, *i.e.,* the angular information, while the rings encode the eccentricity, *i.e.,* the radial information. Such stimuli are presented in figure 5.5. Original results of retinotopic mapping with fMRI can be found in [64, 65, 195].

In standard fMRI protocols, the stimuli presented to the subjects are rotating wedges and expanding rings. The measurements obtained with the rotating wedges provide a polarity map (cf. figure 5.6(a)) while the expanding rings provide an eccentricity map (cf. figure 5.6(b)). A position in the visual field is associated to a cortical position by intersecting the information from the polarity map and the eccentricity map.



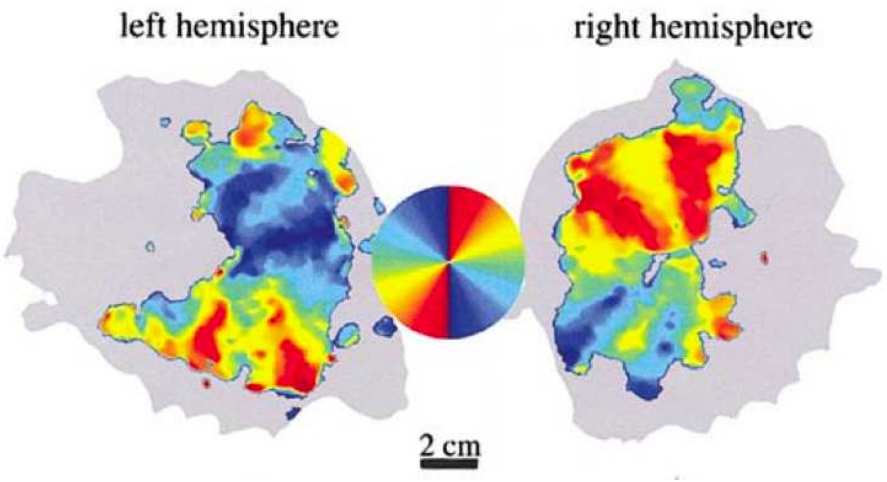
(a) Rings (b) Wedges

Figure 5.5: Rings and wedges visual stimuli used for retinotopic mapping with fMRI (Adapted from [223]).

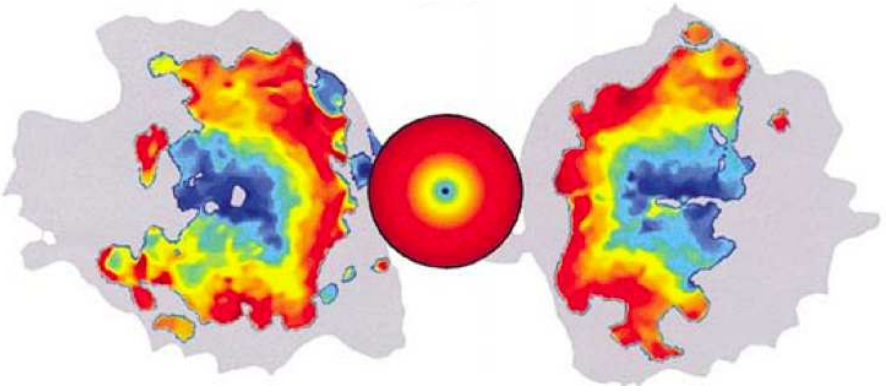
Once the polarity map and the eccentricity map have been computed, the visual areas can be delineated. For example the border between V1 and V2d is given by the lower meridian and the border between V1 and V2v is given by upper meridian. Delineation results are presented in figure 5.7.

The conclusion of this brief presentation of retinotopic mapping with fMRI, is that the spatial resolution of fMRI enables the precise delineation of visual areas like V1, V2 and V3. However, due the low temporal resolution of fMRI, dynamical information remains unaccessible. Retinotopic mapping with M/EEG finds its principal motivation in the measurements of such dynamics.

5.3 Source localization with M/EEG in the visual cortex: previous studies



(a) Orientation, *i.e.,* polarity, map obtained by stimulation using the wedges in figure 5.5(b).



(b) Eccentricity map obtained by stimulation using the rings in figure 5.5(a).

Figure 5.6: Polarity, *i.e.,* orientation, map and eccentricity map obtained by fMRI. Cortical maps are flattened for 2D representation (Adapted from [223]).

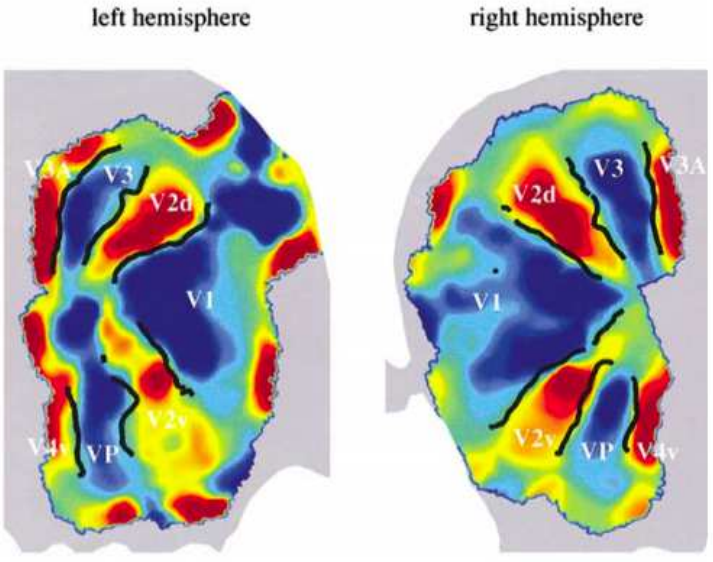


Figure 5.7: Visual areas delineated by fMRI (Adapted from [223]).

We now present a review of important previous studies involving M/EEG recordings for non invasive investigations in the human visual cortex. This presentation is restricted to contributions that achieve source localization.

The literature on the topic demonstrates that various experimental and methodological strategies are possible. Designing a strategy implies to choose:

- the patterns used for stimulation,
- the experimental protocol (duration of presentation of the stimuli, number of repetitions, *etc*.),
- the method for signal extraction (averaging, spectral analysis, *etc*.),
- forward modeling (Spherical head model, BEM, *etc*.),
- inverse modeling (dipole fitting, beamforming, distributed sources, *etc*.).

We now discuss the different possible choices.

Stimulation pattern

Neural response is generated by pattern onset stimuli. This response is however different depending wether the stimuli is black and white or colored, if it is highly contrasted, if it is oriented, *etc*. A good pattern should evoke a response in the visual cortex with a high SNR in order to provide the best source localization results. Following fMRI protocols, most previous M/EEG studies use black and white checkerboard patterns [69, 110, 151]. The patterns are displayed on a gray background to be well contrasted. However, contrary to fMRI, the patterns are presented in a portion of the visual field as illustrated in figure 5.8 with a circular checkerboard pattern. This is done in order to avoid crosstalks between neural current generators. With expanding rings, currents produced by two sources on the walls of the left and right cuneus (cf. figure 5.2) could cancel their effect on the sensors. Source localization would then become impossible.

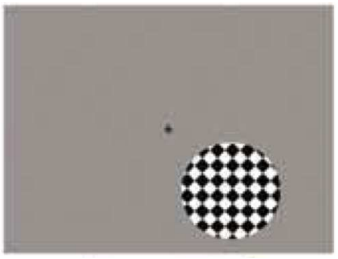


Figure 5.8: Circular checkerboard pattern used for visual stimulation in [151].

Experimental protocol and signal extraction

The experimental protocol details the way the pattern is presented but also what is asked of the subject during the experiment. In requires to choose the duration of the pre-stimulation and stimulation periods and also the inter-stimulation interval (ISI), usually set to be random to limit habituation and anticipation of the subject. The pre-stimulation period is also known as the baseline period. What drives the choice of the protocol is the way the signal of interest is extracted from the measurements. For this purpose, two approaches exist.

The first, and most classical, consists in averaging the measurements of multiple recordings and using as signal of interest the averaged response at a particular time instant. This

instant classically corresponds to a latency peak. Visual evoked potential (VEP) are elicited by pattern onset stimuli. Like for all evoked potentials, the waveform of VEP exhibits characteristic peaks. For VEP, some of these peaks are known as the C1 (a.k.a. N75), the P1 (a.k.a. P100) or the N1 (a.k.a. N145). A typical EEG VEP waveform is presented in figure 5.9. In [110, 214], the authors estimate source amplitudes using the C1. The C1 occurs approximately 70 ms after pattern onset. According to previous fMRI studies and the results described in [110, 214], the neural generators of the C1 are located in V1. With such protocols, the pattern is presented a few hundred times and the pre-stimulation and stimulation periods are respectively about 100 ms and 500 ms. The main problem of such approaches is that the latency peak is not stable across subjects. This implies that the latency of interest on each subject needs to be set manually.

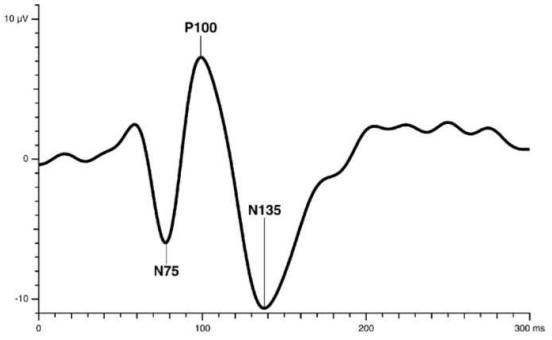


Figure 5.9: A normal pattern reversal VEP measured in EEG (Adapted from [216])

The second consists is extracting the signal of interest from the spectrum of the measured time series. This alternative to VEP is called steady-state visual evoked potential (SSVEP). Rather than displaying multiple time locked pattern onsets, the pattern is flashed on the screen with a known frequency. The stimulus is said to be tagged in frequency. We speak of "frequency tagging". While protocols based on standard VEP focus on the *transient period* just after pattern onset, the experimental paradigms based on SSVEP exploit the *stationarity* in the measured time series. The advantage of working with stationary time series is that standard signal processing tools like Fourier analysis are well adapted to extract the signal of interest. Contrary to the transient VEP signal, the SSVEP signal is easily quantified in the frequency domain and can be rapidly extracted from background noise. A possible drawback is that stimulation periods need to be sufficiently long to let the neural generators enter in a stationary mode. Also, SSVEP are not particularly pleasant for the subject.

SSVEP were previously measured on anaesthetized cats [183] using invasive techniques called multi-unit activity (MUA) and local field potentials (LFP). It was observed that about 300 ms after flicker onset, responses stabilized and exhibited a highly regular oscillatory pattern precisely locked to the stimulus. The stationary period started after 300 ms. This duration for the transient period was confirmed in humans with EEG in [187]. Using such frequency tagged stimuli, Regan [187] observed also that using a "on-off" stimulation paradigm or a reversing pattern does not produce neural activations with the same Fourier spectrum. Pattern reversal stimulation produced a peak in the Fourier spectrum at the double of stimulation frequency [69, 187, 191]. This can be explained by the observation that pattern reversal produce a change of contrast in the visual field twice per period, *i.e.,* twice per full cycle of stimulation. The experimental results adapted from [69] presented in figure 5.10 confirm this observation. The reversing pattern was presented for a long period of time, *i.e.,* 14 s, for 4 different frequencies (2 Hz, 4 Hz, 8 Hz and 21 Hz). In [69] the stimulus consists both of square wave modulated checkerboards or sinusoidally modulated checkerboards, while in [191] the

checkerboards are only modulated by a sine wave. Fawcett *et al*. [69] report that the results with both types of modulations are very similar.

Inverse modeling

As described in chapter 3, the inverse problem can be solved with three main categories of methods: dipole fitting, beamforming and image-based methods with distributed source models. Previous studies on the human vision with M/EEG make use of all three of them.

The most common is dipole fitting. In [214], a small number of dipoles are positioned in the visual areas localized a priori with fMRI and their amplitudes are then estimated. Note that such a procedure is made possible by the fMRI since, as reminded in chapter 3, parametric dipole fitting is a non convex problem for more than one dipole. The positions and orientations of multiple dipoles could not be robustly estimated with pure M/EEG. In [82], the authors use single dipole fitting with MEG data to localize a generator in V1. In [191], Di russo *et al*. use dipole fit with "proximity seeding" and fMRI localizers.

In [69], Fawcett *et al*. investigated with MEG the neural response of V1 to frequency tagged stimuli with a beamforming technique called Synthetic Aperture Magnetometry (SAM). The beamforming method provided the amplitude over time of a manually positioned source in V1. The position of the source was set according to the known position of the pattern in the visual field. The computed time frequency decomposition presented in [69] is reproduced in figure 5.10.

Alternative to dipole fitting approaches and beamforming techniques are imaging methods with distributed source models. In [110] is presented an EEG study with ℓ_2 source estimates. Localization results are compared to fMRI data and it is observed that with the proposed data processing pipeline the spatial resolution of EEG is approximately 3° of visual angle. This study is however limited to visual stimuli displayed on the horizontal meridian. No retinotopic mapping is performed per se. In [197], Sharon *et al*. demonstrate that combining MEG and EEG measurements can improve localisation results compared to pure MEG inverse modeling. The inverse solver is dSPM (cf. chapter 3) which is a noise normalized solver based on an ℓ_2 prior. In [197], no retinotopic maps are presented either.

In 2002, Moradi *et al*. [151] have proposed a quantitative comparison between localization results with MEG and fMRI. The inverse problem is solved on a volumetric grid of distributed dipolar sources and source amplitudes are estimated with an inverse method called Magnetic Field Tomography (MFT) [3]. Results obtained by Moradi *et al*. are presented in figure 5.11. Moradi *et al*. argue that distributed source models are more adapted than dipole fitting since the neural activations evoked by stimulus onsets are very likely to come also from extrastriate regions. Sources amplitudes are estimated at various peak latencies starting around 45 ms after visual presentation of the stimulus. Note that in this contribution, the sources do not lie on a triangular mesh but on a 3D grid. With such a method, like with fMRI data, obtaining retinotopic mapping requires to interpolate activations to the cortical surface.

The different studies detailed above demonstrate the ability of M/EEG to localize with a good precision neural activations in the intricate occipital region around the calcarine sulcus. This observation was also confirmed by the simulation study we published in [2]. However none of these contributions provides a complete retinotopic mapping of V1. Dipole fitting and beamforming approaches are certainly not adapted for such a purpose that can only be achieved with a distributed source model and an image-based inverse solver. The methodologies developed in [110] and in [197] are probably the best attempt towards retinotopic mapping with MEG and therefore the closest related work to the contribution presented in this chapter. These methods however work with VEP and therefore suffer from various drawbacks like the problem of variable peak latencies between subjects. Their experimental results are also limited to the horizontal meridians in [110] and the four quadrants in [197].

The use of frequency tagged stimuli in [69, 151] demonstrates the ability of M/EEG devices

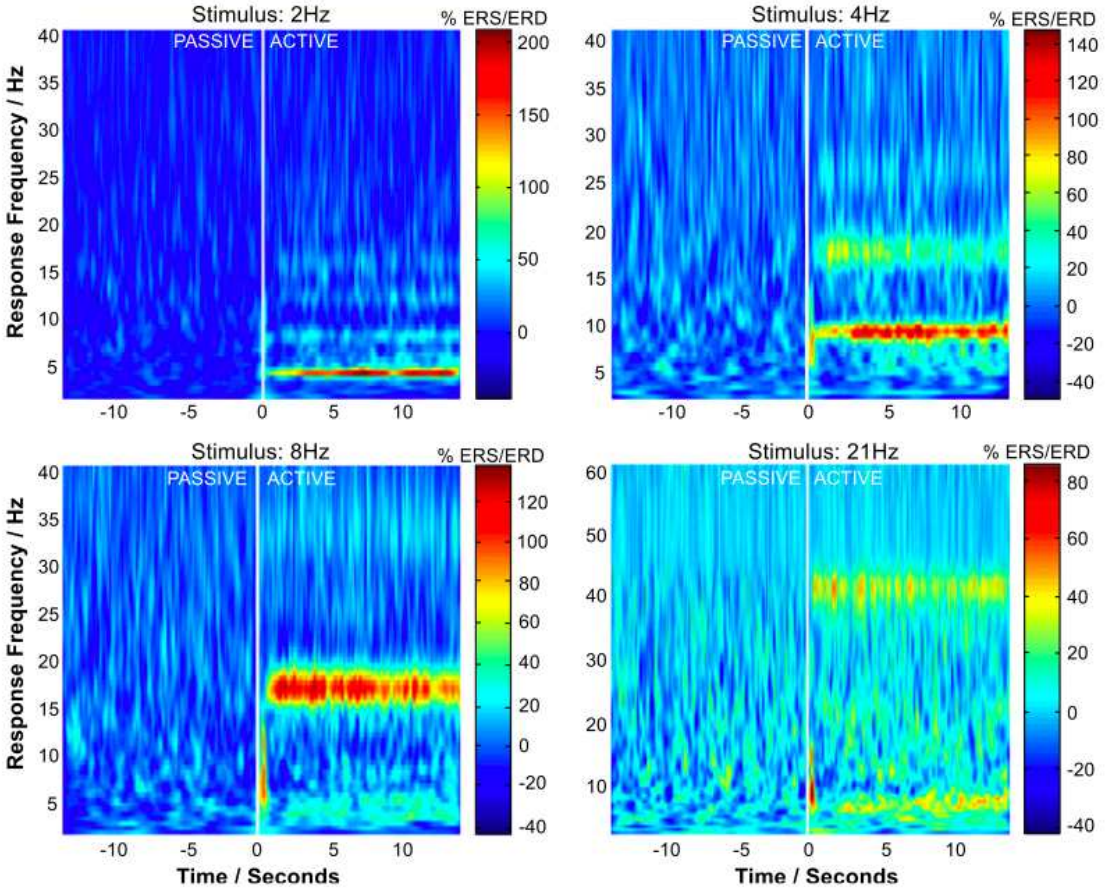


Figure 5.10: Time-frequency plots obtained using a checkerboard pattern flickering at various frequencies in the bottom right quadrant of the screen. Time-frequency plots are calculated using Morlet wavelet analysis from a voxel in the left (that is, contralateral) medial visual cortex and averaged across multiple subjects to improve the SNR. The cortical harmonics of the stimulus frequency can be clearly seen in the active phase, together with an onset response shortly after the stimulus onset. The colour bar on the right of each figure shows the event-related synchronization (ERS) and event related desynchronization (ERD) scale used, expressed as a percentage of change from the baseline. (Adapted from [69])

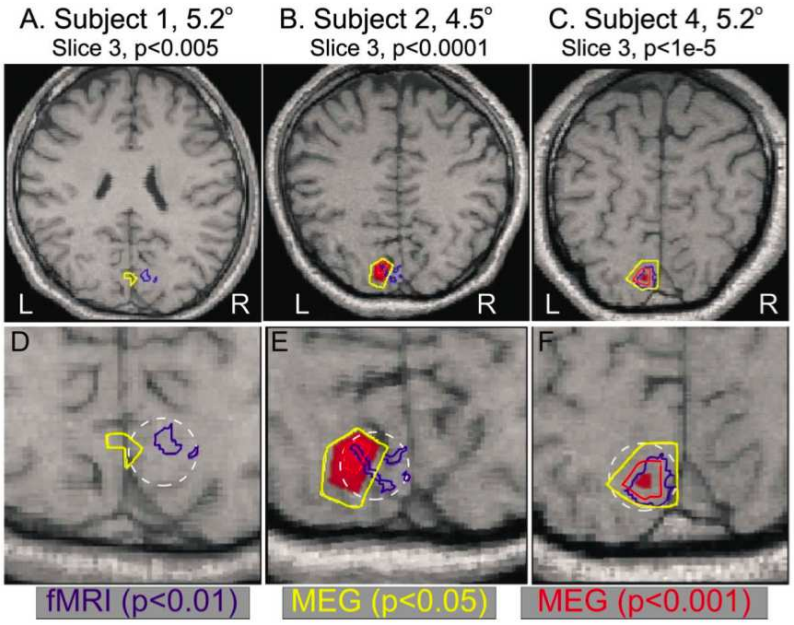


Figure 5.11: Localization results obtained by Moradi in [151] with fMRI and MEG. Results are displayed on a 2D slice going through the calcarine fissure. In blue are the results obtained by fMRI (p¡0.01) while in yellow and red are the active regions estimated with MEG (p¡0.01 and p¡0.001). (Adapted from [151])

to capture with a good SNR the signal evoked by a flickering pattern. Contrary to variable peak latencies observed with VEP on multiple subjects, the frequency tuning of the neural response evoked by SSVEP appears to be much more stable across subjects.

For these reasons, the protocol proposed in the following section is based on SSVEP and the inverse problem is solved with a distributed source model. Besides the protocol, our contribution is on the efficient processing of the data, both for the signal extraction and the inverse modeling.

5.4 MEG EXPERIMENTAL DESIGN

5.4.1 Stimulus design

The primary objective of the design of the stimulation pattern was to produce a good signal to noise ratio (SNR) in the MEG measurements. Considering what has been said earlier on the functional properties of the simple cells in the visual cortex, a good SNR can be obtained with highly contrasted patterns presenting multiple orientations. By multiplying the number of orientations, more simple cells get activated and the global amplitude of the evoked response of the brain increases.

The stimuli that have been designed are presented in figure 5.12. In order to obtain a mapping of V1, the “star” like patterns have been placed at multiple positions of the visual field. Each position corresponds to an experimental condition. There are two types of experimental conditions: the quadrants and the meridians. Each quadrant contains a big single “star”. Along the meridians 4 “stars” whose sizes increase in accordance to the cortical magnification factor in V1 [52] are displayed.

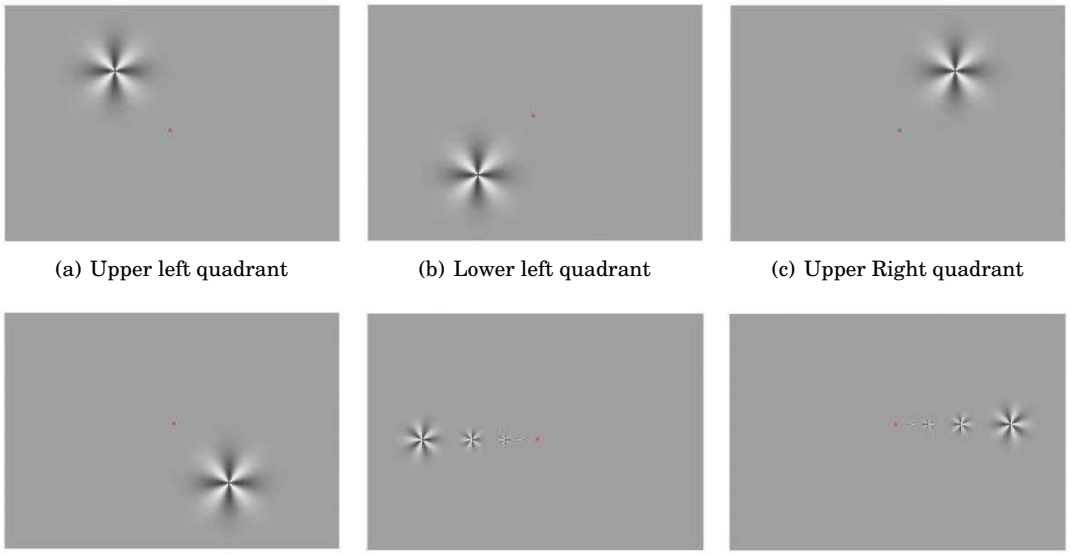
Stimuli were back-projected using a video projector (60 Hz refresh rate) on a translucent screen located away 90 cm from the subjects. The Michelson contrast of the displayed patterns was 96%. They were presented against a grey background (103 cd/m^2; see figure 5.12).

The Michelson contrast is defined by:

$$C_{\text{Michelson}} = \frac{L_{\max} - L_{\min}}{L_{\max} + L_{\min}}$$

where $L_{\min}$ and $L_{\max}$ are respectively the minimum and the maximum luminances (cd/m^2) measured on the pattern.

In each quadrant, the pattern was $2.66°$ wide with an eccentricity of $8.85°$. For the meridians the 4 patterns had varying sizes (Width: $0.22°$, $0.35°$, $0.71°$, $1.33°$) and were presented at 4 different eccentricities ($0.62°$, $1.52°$, $3.3°$, $6.64°$ respectively).



(a) Upper left quadrant (b) Lower left quadrant (c) Upper Right quadrant

(d) Lower right quadrant (e) Left meridian (f) Right meridian

Figure 5.12: Stimuli displayed for retinotopic mapping with MEG. Each position corresponds to an experimental condition. The subject is asked to look at the colored fixation point at the center of the screen. The flickering patterns were displayed for 6.5 s at 7.5 Hz or 10 Hz in each trial evoking a steady-state visual evoked potential (SSVEP).

5.4.2 Protocol design

The primary motivation for using SSVEP and frequency tagged stimuli is that for long enough stimulation periods the visual system can be considered to be *stationary*. The complex dynamic phenomena that occur during the transient period just after a stimulus onset are not considered. Since a SSVEP can be completely described in terms of the amplitude and phase of each frequency component it can be quantified more unequivocally than an averaged transient evoked potential.

The choice of the stimulation frequency

In [69], Fawcett *et al*. presented a reversing pattern tagged with 4 different frequencies. The time frequency maps presented in figure 5.10 demonstrate that the increase in synchronization during stimulation, i.e, active period, exceeds 100% at 2, 4 and 8 Hz. At 21 Hz the steady-state response is clearly reduced. Due to the reversing pattern, the actual frequency

observed corresponds to the second harmonic. Similar experimental facts are reported by Pastor *et al*. in [173] with EEG measurements and a flickering "on-off" pattern (no contrast reversing). In order to estimate the best stimulation frequency for brain computer interface (BCI) based on SSVEP, the authors measured the amplitude of the signal on the occipital electrodes as a function of the stimulation frequency. As presented in figure 5.13, they observed that the amplitude reached a maximum at 15 Hz and then fell with a plateau up to 27 Hz, declining at higher frequencies.

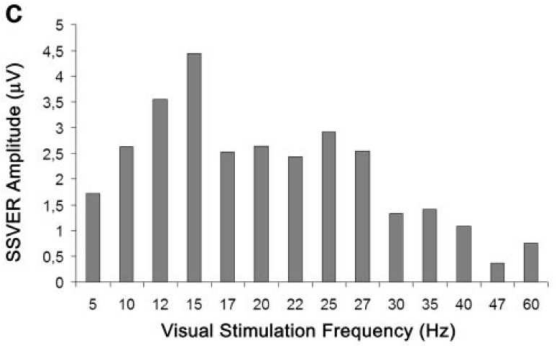


Figure 5.13: Average of the mean values of the amplitude of the FFT fundamental frequency of the SSVEP recorded on three occipital EEG electrodes at the different stimulation frequencies. The amplitude of the occipital neural response, expressed in microvolts, reached a maximum at 15 Hz and then fell with a plateau up to 27 Hz, declining at higher frequencies. (Adapted from [173])

These experimental observations indicate that the best frequency for a stimulation with reversing pattern is between 5 and 10 Hz. Our hypothesis explaining the dependence of the SSVEP amplitude on the stimulation frequency is that the retina is the limiting step of the visual processing pipeline. This is confirmed by the experimental results from [196] reproduced in figure 5.14. By measuring the response of ganglion cells in the retina of an anesthetized cat while it was presented a sinusoidally modulated pattern, Shapley and Victor concluded in 1978 that the retina is tuned to attenuate less the temporal frequencies between 10 and 20 Hz. These experimental results agree with the conclusion obtained by Pastor *et al*. and Fawcett *et al*. with M/EEG measurements on humans.

Taking this into account in our experimental paradigm, stimuli were displayed at a frequency of 7.5 Hz and 10 Hz. For experimental reasons, the stimulation frequency were constrained by the 60 Hz refresh rate of the screen. During the stimulation period, with 7.5 Hz, the pattern was displayed during 4 successive frames before alternating with the reversed pattern during the next 4 frames. At 10 Hz, the pattern was displayed during 3 successive frames before alternating with the reversed pattern during the next 3 frames.

Pre-stimulation and stimulation periods

Stimulation consists of multiple repetitions of alternating pre-stimulation and stimulation periods. A cycle of pre-stimulation and stimulation period is called a trial.

A trial started with the display of a colored fixation disk at the center of the screen during 600 ms. It was followed by one of the patterns displayed in figure 5.12 flickering in counter-phase at 7.5 Hz or 10 Hz. Two successive trials were separated by a random inter stimulus interval (ISI) of about 1.5 s (cf. figure 5.15).

We call "run" the successive presentation of multiple trials. One run was done for the quadrants and one for the meridians. In each run, the conditions were randomly presented. Each run contained 15 trials for each condition, *i.e.,* position in the visual field. For example

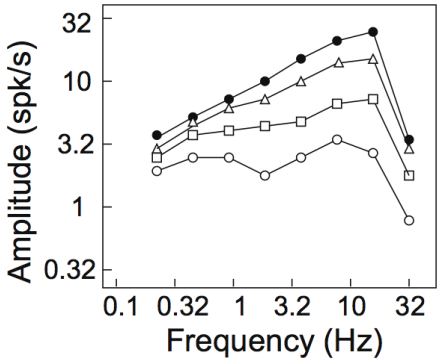


Figure 5.14: Amplitude of response of cat ON-center X ganglion cell, reproduced from Shapley and Victor (1978). Stimulus consists of 4 sinusoids with different contrasts. It reveals that the frequencies between 10 and 20 Hz are the less attenuated by the retina. (Adapted from [228] and original data from [196])

the run for the quadrants had 60 trials with 15 trials for each quadrant.

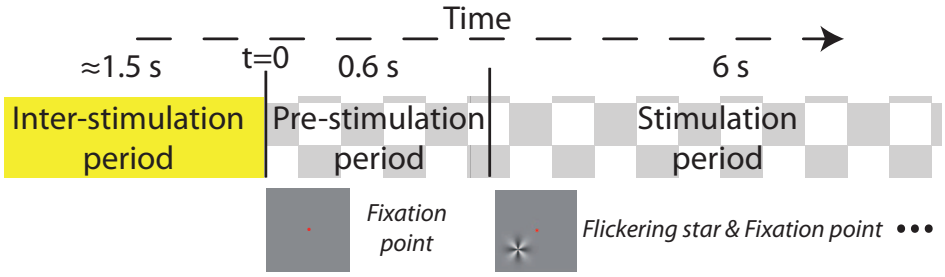


Figure 5.15: A trial in the protocol for retinotopic mapping with MEG. Each trial is composed of an inter trial period also called ISI (inter stimulus interval). This period has a random length close to 1.5 s. At t=0 a fixation point appears and 600 ms later the stimulation starts.

During a trial, the fixation disk randomly changed color every second. To maintain the attention of the subjects focussed and minimize eye movements, they were asked to report on which color appeared more often at the end of the run.

5.5 Mapping V1 with MEG

5.5.1 Data exploration

In order to confirm the presence of energy at the frequencies of stimulation on the data, periodograms and spectrograms were computed on the signals measured by each MEG sensor.

The periodograms provide estimates of the power spectral density (PSD) of the signal during the steady-state period. The stimulation started 0.6 s after the fixed point appeared and the transient regime was estimated around 0.4 s. The beginning of the steady-state period was therefore set to $t = 1\ s$.

The periodogram can simply be computed with a standard FFT or in a more efficient

way using a smooth periodogram technique introduced in [225]. This technique commonly referred to as a multitaper method, consists in computing the PSD on different portions of the data and to average all the results in order to reduce the variance of the PSD estimation. Each portion of the data is tapered by a time domain window. This is illustrated in figure 5.16, where 6 tapers with a 50% overlap are extracted from a single-trial measurement. For each of the 6 signal windows, the PSD is computed. The 6 results are then averaged to provide the periodogram. By doing so the periodogram is smoothed and the PSD estimate is biased. The variance of the estimate is however divided by the number of windows which makes the PSD estimate more accurate. It corresponds to the classical trade-off in statistical estimation between bias and variance.

To illustrate this, PSD estimates computed on the grand average data during one condition are represented in figure 5.17. It can be observed that using small windows lead to a smoothed version of the periodogram. This procedure reduces the estimation bias called *spectral leakage*. When estimating the PSD of finite-length signals or finite-length segments of infinite signals, it happens that some energy *leaks* out of the original signal spectrum into other frequencies. In the present case where the stimulation frequency produces a peak at 15 Hz, it happens that part of the signal of interests leaks to the frequency bins near 15 Hz. The frequency bins are fixed by the FFT depending on the length on the input signal. Therefore, they may not include a bin exactly at 15 Hz.

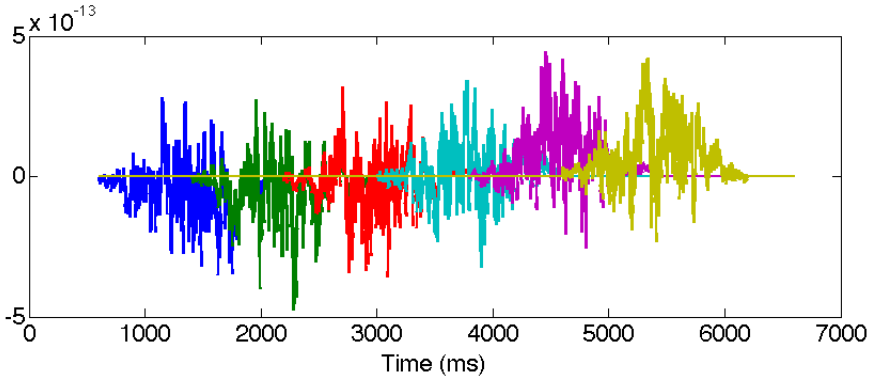
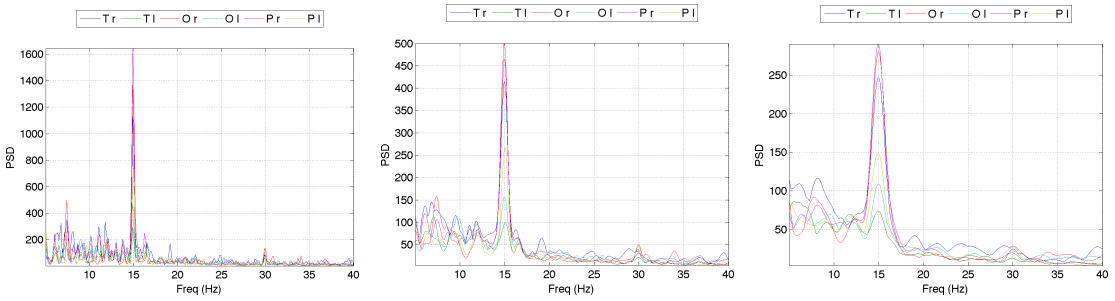


Figure 5.16: Multi-taper example on a single-trial MEG measurement



(a) 1 window of 5 s (b) 5 windows of 1.6 s (c) 12 windows of 0.8 s

Figure 5.17: Multi-taper periodogram obtained with 3 different sizes of windows. Each curve corresponds to one sensor on the different groups of MEG sensors (OR: occipital right, OL: occipital left, PR: parietal right, PL: parietal left, TR: temporal right, TL: temporal left). The periodogram is estimated on the averaged data for one subject during stimulation of the lower left quadrant at 7.5 Hz. One can observe the smoothing effect of the multitapering.

Once the PSD is estimated, the topography at the frequency of interest can be observed.

In figure 5.18, the estimated PSD at 15 Hz (or at the closest frequency bin) is displayed. The topography represents an energy and is therefore positive. A *hot spot* on the occipital sensors can be observed which confirms the spatial localization of the signal.

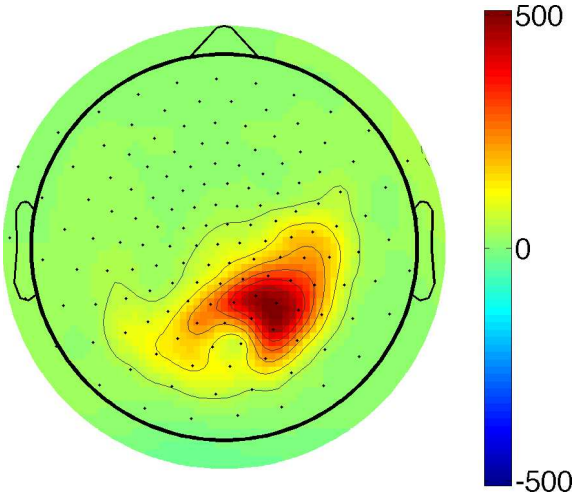


Figure 5.18: PSD at 15 Hz represented on the sensors. A clear hot spot on the occipital region confirms the presence of a source responding at 15 Hz in the occipital cortex.

Another convient way for exploring the spectral content of the data, consists in computing time-frequency (TF) maps. Using filters banks localized in time and frequency, such representations describe the location of energy in a time-frequency plot. Common filters used in M/EEG are Gabor and Morlet filters. We refer the reader to Appendix C for more details on how the Gabor filters used in this work were designed.

An interest of the TF representations is that they enable a comparison between the pre-stimulation and the stimulation periods. Let us denote $TF(t, f)$ the PSD estimated at time t and frequency f. Let $TF_{base}(t, f)$ the restriction of $TF(t, f)$ to the pre-stimulation period ($0 < t < 0.6$ s) and $\overline{TF_{base}(f)}$ the mean over t of $TF_{base}(t, f)$. Let us call the Event Related Synchronization / Event Related Desynchronization (ERS/ERD) coefficient the quantity:

$$\mathrm{ERS/ERD}(t, f) = \frac{TF(t, f) - \overline{TF_{base}(f)}}{\overline{TF_{base}(f)}} \quad .$$

If no difference is present between pre-stimulation and stimulation periods $\mathrm{ERD/ERS}(t, f)$ is equal to 0. If $\mathrm{ERD/ERS}(t, f)$ is equal to 10, it means that the energy during stimulation is 10 times bigger than during the pre-stimulation, a.k.a., the baseline period. To have a representation that scales to possibly large values of $\mathrm{ERD/ERS}(t, f)$, we represented in the TF plot the quantity:

$$\mathrm{sign}(\mathrm{ERD/ERS}(t, f)) \log(1 + |\mathrm{ERD/ERS}(t, f)|) \quad ,$$

where $\mathrm{sign}(x)$ stands for the sign of x, *i.e.,* $\mathrm{sign}(x) = x/|x|$ if $x \neq 0$ and 0 otherwise. The function $\log$ stands for the decimal logarithm ($\log(10) = 1$).

In figure 5.19, a sample TF plot is presented. Like in [69], the TF plot was computed on the signal obtained by averaging all the trials corresponding to the condition of interest. We can observe a strong increase of PSD at the harmonics of $f = 7.5$ Hz, especially $2f = 15$ Hz, after the stimulation onset at 0.6 s. The harmonic $4f = 30$ Hz can also be observed. The dashed vertical bar represents the beginning of the stimulation. This observation, done also on different sensors and conditions, led us to the conclusion that the signal of interest could

be extracted from the Fourier coefficients at 15 Hz.

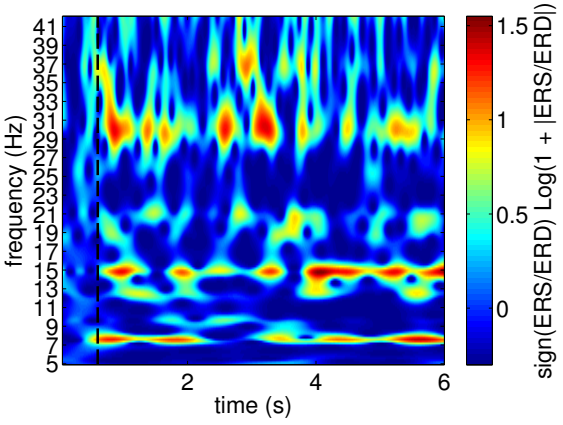


Figure 5.19: Sample time frequency map, a.k.a., spectrogram, estimated on the averaged signal measured on the MLO11 sensor. Stimulation was performed at 7.5 Hz with a pattern positioned on the lower left quadrant. Spectrogram was computed with Gabor filters with $\xi = 15$ (cf. chapter C).

5.5.2 Method

5.5.2.1 How to invert?

In order to localize the current generators responding to the frequency tagged stimulation pattern, the straightforward way consists in inverting the full temporal data in order to obtain a time series per current dipole on the cortex. Then for each dipole the PSD can be computed. The dipole having the highest PSD are the most likely to be active.

Using a linear inverse method based on an ℓ_2 prior like the standard minimum-norm (MN) solver, the source estimates are given by:

$$\mathbf{X} = \mathbf{G}^T(\mathbf{G}^T\mathbf{G} + \lambda\mathbf{I})^{-1}\mathbf{M} \ .$$

Spectral estimation with Fourier analysis is linear. Let us denote $\mathbf{\Phi}$ the dictionary of Fourier atoms such that the Fourier transform of a temporal vector $\mathbf{x}$ can be written in a matrix form $\mathbf{\Phi x}$. The Fourier coefficients of the sources can then be written $\hat{\mathbf{X}} = \mathbf{X}\mathbf{\Phi}^T$. The ith line of $\hat{\mathbf{X}}$ contains the Fourier transform of the temporal activation of the ith dipole on the source space. By writing this in matrix form, it appears that the Fourier transform of the sources can be obtained from the Fourier coefficients estimated on the sensors. We have that $\hat{\mathbf{X}} = \mathbf{G}^T(\mathbf{G}^T\mathbf{G} + \lambda\mathbf{I})^{-1}\hat{\mathbf{M}}$ where $\hat{\mathbf{M}} = \mathbf{M}\mathbf{\Phi}^T$. The practical consequence of this observation is that when using a linear inverse solver the FFT on the sources can be obtained at the price of the computation of the FFT on the sensors. The PSD can then be obtained from the FFT. Rather than inverting the full data, this suggests to invert only the portion of the spectrum that contains the relevant information. In the current study, this information appears to be contained in the Fourier coefficient at 15 Hz.

5.5.2.2 Estimating active regions with permutation tests

The inverse problem gives the current estimates. However, what is really of interest are the locations where these estimates are sufficiently significant to be considered as active. In order to locate these statistically significant activations, non-parametric statistical tests provide a

robust and principled approach.

Basics of non-parametric statistical tests and permutations

H_0 and H_1. Like standard statistical tests, non-parametric tests are used to test the validity of a hypothesis. There two hypotheses: the null hypothesis, denoted H_0, that states that observations result purely from chance, and the alternative hypothesis, denoted H_1, that the observations are influenced by a non random cause. Rejecting H_0 with a certain level of confidence means that the observations are not purely the result of chance.

In the context of brain functional imaging with M/EEG, the null hypothesis H_0 classically states that the stimulation has no effect on the neural activation in a given brain region, while H_1 states that the activation is influenced by the stimulation. When rejecting H_0, the brain region of interest is flagged as active.

Type I and type II errors. Two types of errors can be erroneously produced in statistics: type I and type II. A type I error is made when the null hypothesis is erroneously rejected. It corresponds to a false positive. Type II errors are made when the null hypothesis is erroneously accepted. It corresponds to a false negative. The statistical power of a test is related to type II errors. The more powerful a test is, the less it will produce type II errors. In practice, one wants to use a statistical test with a high power and with a control of type I errors. This control is given by the P-value.

Let us illustrate this with an example in the context of non-parametric tests.

Example. Let us consider two populations of 10 subjects each. The two populations, denoted A and B, have been assigned to 2 different treatment conditions (conditions A and B). If the null hypothesis H_0 is true, no difference between the means of condition A and B will be found at the end of the treatment. Under H_0, *i.e.,* if H_0 is true, label A is not different from label B, hence they can be exchanged without affecting the difference between means A and B. In the same way, the 10 observations can be switched from A to B. Let us now systematically rearrange the 20 observations by permuting the labels of the observations and compute the difference between mean A and B. There are $2^{10} = 1024$ possible permutations of the labels, leading to 1024 values of difference of the means between A and B under the null hypothesis. We can now answer the following (inferential) question: Under H_0, what is the probability to obtain the difference in the means that we observed in the experimental data? If this quantity is smaller than 5%, we say that the hypothesis H_0 is rejected with a P-value of 0.05.

Let us rewrite this with simple equations. Let us denote $\mathbf{a} = (a_l)_l \in \mathbb{R}^{10}$ (resp. $\mathbf{b} = (b_l)_l \in \mathbb{R}^{10}$) the vector of measurements observed in population A (resp. B). The difference of the means obtained with experimental data is:

$$T_{exp} = \frac{1}{10}\sum_l (a_l - b_l) = \frac{1}{10}(\mathbf{a}-\mathbf{b})^T \cdot \mathbf{1} \ ,$$

where $\mathbf{1}$ is a vector in $\mathbb{R}^{10}$ filled with ones. It corresponds to the value of the statistic, denoted T, with the experimental data. We denote it T_{exp}. In order to compute a "permuted" version of this statistic one can simply replace the vector $\mathbf{1}$ with a vector filled randomly with 1 and -1. For a given permutation indexed by n we denote $\mathbf{p}_n$ this random vector. We compute a value of the statistic under H_0 with:

$$T_n = \frac{1}{10}(\mathbf{a}-\mathbf{b})^T \mathbf{p}_n \ .$$

Let us denote $\mathbf{P} \in \mathbb{R}^{10\times 1024}$ a matrix containing as columns all the possible vectors $\mathbf{p}_n$. The distribution under H_0 is obtained by:

$$(T_n)_n = \frac{1}{10}(\mathbf{a}-\mathbf{b})^T\mathbf{P} \in \mathbb{R}^{1024} \ .$$

The hypothesis H_0 is rejected with a P-value of 0.05 if:

$$\frac{\#\{n/T_n \geq T_{exp}\}}{1024} \leq 0.05 \ .$$

General permutation test procedure.

1. Select a test statistic which measures the differences between conditions (here difference of the means).
2. Compute the test statistic for the original condition labeling (T_{exp}).
3. For each resampling k, randomly rearrange the condition labels and compute the test statistic T_k for the permuted data and add it to the null distribution.
4. Repeat step 3 until a predefined number of resamplings has been performed (or all resamplings if it is tractable).
5. Compare the null distribution of the test statistic to the original data.
6. Accept or reject the null hypothesis based on the proportion of permuted test statistics greater than or equal to the original.

Remarks. Using as statistics the difference of the means has the advantage that the procedure can be described with basic linear algebra formulas. This formulation with matrices also has the advantage to provide a straightforward way to implement it and to benefit from efficient linear algebra software packages like (BLAS/LAPACK or the Intel MKL).

In practice, one often has more than 10 observations per condition. Therefore testing all the possible permutations is not always tractable. To circumvent this problem only a random sample of all data permutations is selected. By doing so, the test loses of his statistical power and the P-value is only an approximation. It can be observed that the P-value cannot be smaller than 1/1024 or more generally $1/N$, where N is the number of random permutations.

Non-parametric tests are very flexible compared to standard parametric tests. Any statistic, such as a T-test, can be used depending of the application. Also, no assumption, such as Gaussianity, on the statistical distribution of the random samples is required. More importantly, they offer a very easy way to correct for the problem of multiple comparisons.

Multiple comparisons. When running many times the same test, assuming the tests are independent, the probability that one test will reject H_0 increases. When running 100 tests with independent samples with a P-value of 0.01 this probability is equal to 1! The null hypothesis will be rejected once almost surely. Therefore, when running multiple tests, it is important to control for example the FWER (Familywise error rate). The FWER is the probability of making one or more type I errors among all the hypotheses when performing multiple tests.

Remark. The False Discovery Rate (FDR), often used in neuroimaging [85], is a different way of controlling Type I errors when running multiple comparisons. FDR controls the expected proportion of incorrectly rejected null hypotheses. If the FDR is controlled with a P-value of 0.05, it means that among 100 rejected null hypotheses, 5 are expected to be false positives. With a control of the FDR, one rejects null hypotheses more “easily” than when controlling

the FWER. The control of the FDR is sometimes said to be less conservative than the FWER.

With non-parametric tests, control of the FWER can be done using what is called the statistic of the "max". Let us assume that a test is performed at multiple brain locations, indexed by i.

$$\begin{aligned} P(FWER) &= P(\cup_i \{T^i \geq u\} | H_0) \text{ (Prob. any position exceeds the threshold } u) \\ &= P(\max_i T^i \geq u | H_0) \text{ (Prob. max position exceeds the threshold)} \\ &= 1 - F_{\max T | H_0}(u) \text{ (1-cumulative density function of max position)} \\ &= 1 - (1-\alpha) = \alpha \end{aligned} \tag{5.1}$$

This means that controlling the probability that the maximum statistic over the brain exceeds a threshold u under H_0 provides a control of the FWER with the same probability α.

Let us come back to our example. Suppose now that we run the test of the two treatment conditions A and B, $I > 1$ times. The measured results are stored in two matrices $\mathbf{A} \in \mathbb{R}^{10 \times I}$ and $\mathbf{B} \in \mathbb{R}^{10 \times I}$. The experimental values of the statistic are given by:

$$(T^i_{exp})_i = \frac{1}{10}(\mathbf{A} - \mathbf{B})^T \cdot \mathbf{1} \in \mathbb{R}^I \ .$$

Under H_0 the distribution of the maximum is given by:

$$(T^{max}_n)_n = \max\left(\text{abs}\left(\frac{1}{10}(\mathbf{A} - \mathbf{B})^T \mathbf{P}\right)\right) \in \mathbb{R}^{1 \times 1024} \ ,$$

where the function $\max$ computes the maximum value of every column of its input and abs computes the modulus of each coefficient of its input. The hypothesis H_0 is rejected at position i with a P-value of 0.05 if:

$$\frac{\#\{n / T^{max}_n \geq T^i_{exp}\}}{1024} \leq 0.05 \ .$$

Sample histograms of $(T^i_{exp})_i$ and $(T^{max}_n)_n$ are presented in figure 5.20. Data are extracted from the retinotopy MEG dataset for which I is between 15000 and 40000 depending on the number of dipoles considered on the cortical mesh.

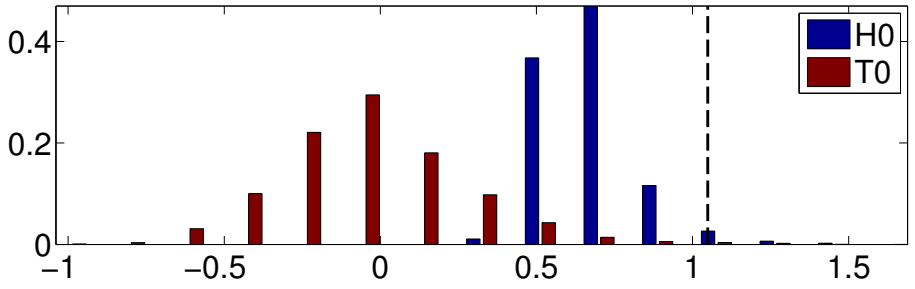


Figure 5.20: Example of histograms of $(T^i_{exp})_i$, denoted T0, and histogram of the $(T^{max}_n)_n$, denoted H0. The dashed vertical line represents the threshold of the statistic with a P-value of 0.05. 5% of the $(T^{max}_n)_n$ are above this line. The null hypothesis is rejected at all the positions i such that T^i_{exp} is above this threshold.

Permutations tests with M/EEG data. In neuroimaging, and particularly with M/EEG data, what is of interest is to know where and when a particular brain region is activated by the stimulation. The control condition is then typically extracted from the measurements recorded in the pre-stimulation period, a.k.a., the baseline period. Permutation tests were introduced in the field of M/EEG in [168, 169].

5.5.2.3 The mapping procedure

The data exploration on the sensors in section 5.5.1 confirmed the presence of significant information at 15 Hz. In section 5.5.2.1, it was explained how to restrict the computation of the inverse problem to a portion of Fourier spectrum. And in section 5.5.2.2, basics of non-parametric statistical tests have been presented. The mapping procedure that we now present is based on all these remarks and results.

The frequency of interest is 15 Hz. Therefore, rather than computing an FFT to get the Fourier coefficients at 15 Hz, or at the closest frequency bin, a simple correlation of the signal with a complex sinusoid tuned at 15 Hz can be used. Such a sinusoid at a frequency f_0 is defined by

$$\phi_{f_0}(t) = \exp(2i\pi f_0 t) \ .$$

The discretized version of ϕ_{f_0} is a vector denoted $\boldsymbol{\phi}_{f_0}$. By computing the correlation of the measurements in each trial, indexed by l, with $\boldsymbol{\phi}_{f_0}$, we obtain a complex valued Fourier coefficient for each sensor in each trial. Let us denote $\mathbf{M}_l$ the measurements for trial l. The coefficients for this trial are given by $\hat{\mathbf{m}}^l_{f_0} = \mathbf{M}_l \boldsymbol{\phi}_{f_0} \in \mathbb{R}^{d_m}$. By concatenating all the $\hat{\mathbf{m}}^l_{f_0}$ for all the d_l trials, we get a matrix of Fourier coefficients, $\hat{\mathbf{M}}_{f_0} \in \mathbb{R}^{d_m \times d_l}$. For the sake of simplicity, we will omit from now on the index f_0 for the Fourier coefficients.

The data that we propose to invert is $\hat{\mathbf{M}}$. Using an ℓ_2 prior, the Fourier coefficients in the source space are given by:

$$\hat{\mathbf{X}} = \mathbf{G}^T(\mathbf{G}^T\mathbf{G} + \lambda \mathbf{I})^{-1}\hat{\mathbf{M}} \in \mathbb{R}^{d_x \times d_l} \ .$$

In order to run statistical tests, the Fourier coefficients need to be estimated under two conditions, here on the stimulation period and on the baseline period. Let us denote $\hat{\mathbf{X}}^{stim}$ and $\hat{\mathbf{X}}^{base}$ the two sets of coefficients. If we were to consider the difference of the means as statistic the distribution under H_0 would be given by:

$$(T_k^{max})_k = \max\left(\text{abs}\left(\frac{1}{d_l}(\hat{\mathbf{X}}^{stim} - \hat{\mathbf{X}}^{base})^T\mathbf{P}\right)\right) \in \mathbb{R}^{1\times N} \ .$$

The matrix $\mathbf{P} \in \mathbb{R}^{d_l \times N}$ is the permutation matrix filled with 1 and -1.

In order to compensate for depth bias, it is classical with M/EEG to normalize the reconstructed currents using an estimate of the variance of the noise. We refer the reader to chapter 3 and particularly to section 3.2.2.3. The estimate of the noise variance is obtained by computing the variance and standard deviation of each row vector in $\hat{\mathbf{X}}^{base}$. We denote this vector of standard deviations by $\boldsymbol{\sigma}^{base} = (\sigma_i^{base})_i \in \mathbb{R}^{d_x}$. The noise normalized versions of $\hat{\mathbf{X}}^{stim}$ and $\hat{\mathbf{X}}^{base}$, denoted $\hat{\mathbf{X}}^{stim}_{nn}$ and $\hat{\mathbf{X}}^{base}_{nn}$, are obtained by dividing each coefficient on line i by the standard deviation σ_i^{base}.

The experimental value of the statistic at position i, T^i_{exp}, is then given by:

$$T_{exp} = \frac{1}{d_l}(\hat{\mathbf{X}}^{stim}_{nn} - \hat{\mathbf{X}}^{base}_{nn})^T \cdot \mathbf{1} \in \mathbb{R}^{d_x} \ , \tag{5.2}$$

and the distribution under H_0 by:

$$(T_k^{max})_k = \max\left(\text{abs}\left(\frac{1}{d_l}(\hat{\mathbf{X}}_{nn}^{stim} - \hat{\mathbf{X}}_{nn}^{base})^T\mathbf{P}\right)\right) \in \mathbb{R}^{1\times N} \ .$$

Again, the hypothesis H_0 is rejected at position i with a P-value of 0.05 if:

$$\frac{\#\{n/T_n^{max} \geq T_{exp}^i\}}{N} \leq 0.05 \ .$$

The vertices on the triangulated source space where H_0 is rejected, are the active vertices.

Computation time. The Fourier coefficients of interest are obtained with a simple matrix multiplication on both the stimulation and the baseline data. The inverse computation with an ℓ_2 prior, is also achieved with a simple matrix multiplication. Finally, thanks to our presentation of the permutation tests procedure using also a matrix formulation in section 5.5.2.2, the full procedure has a very limited computational demand. On a standard computer, the mapping pipeline takes less than a 1 minute. This of course assumes that the forward models have been previously computed.

Testing on the Fourier coefficients vs. on the Power spectral density. The procedure just described works on the Fourier coefficients. An alternative strategy consists in working on the power spectral densities (PSD). Rather than considering T_{exp} as detailed in (5.2), the statistic can be computed with:

$$T_{exp} = \frac{1}{d_l}(|\hat{\mathbf{X}}_{nn}^{stim}|^2 - |\hat{\mathbf{X}}_{nn}^{base}|^2)^T \cdot \mathbf{1} \in \mathbb{R}^{d_x} \ .$$

The matrix $|\hat{\mathbf{X}}_{nn}|^2$ contains the squared modulus of the elements of $\hat{\mathbf{X}}_{nn}$. A possible motivation for using the PSD rather than the Fourier coefficients with the phase information is that it allows to estimate the PSD using a multitaper approach. However, by neglecting the phase, the statistical power of the test is reduced. But on the other hand, the quality of the estimate of the PSD, or equivalently the modulus, is improved. In our study both approaches were investigated.

5.5.3 Mapping results

5.5.3.1 Localization results with ℓ_2 inverse solvers

The first step consists in computing $\hat{\mathbf{X}}^{stim}$ and $\hat{\mathbf{X}}^{base}$. An estimate of the noise can be estimated from $\hat{\mathbf{X}}^{base}$. This leads to the computation of $\hat{\mathbf{X}}_{nn}^{stim}$ and $\hat{\mathbf{X}}_{nn}^{base}$, and finally T_{exp}. A example of T_{exp} map is represented in figure 5.21. This image was obtained with a stimulation in the lower left quadrant of the visual field with a pattern flickering at 7.5 Hz. The quantity T_{exp} was here computed using the Fourier coefficients at 15 Hz (not the PSD).

In order to interpret the values displayed, it can be noticed that the value T_{exp} can be related to a *z-score*. A value of 2 is a value twice bigger than the variance estimated on the baseline. For a Gaussian distribution, a value twice bigger than the standard deviation is relatively significant. However, the active regions were not designated with such a rule of thumb. The map in figure 5.21 is thresholded using the non-parametric statistical procedure detailed above. The resulting thresholded activation map is presented in figure 5.22. The image was obtained with a p-value set to 0.05 and 15000 permutations. The experimental data contained 15 trials which means that we used 15000 out of $2^{15} = 32768$ possible permutations. In order to visually improve the result, we extracted from the thresholded map

the connected component containing the highest activations (cf. figure 5.22(c)). We observe in this result, that the estimated active region stands on the upper bank of the calcarine fissure. According to the position of the stimulation pattern in the lower left quadrant, this localization result agrees perfectly with our knowledge on the organization of the primary visual cortex (cf. section 5.1).

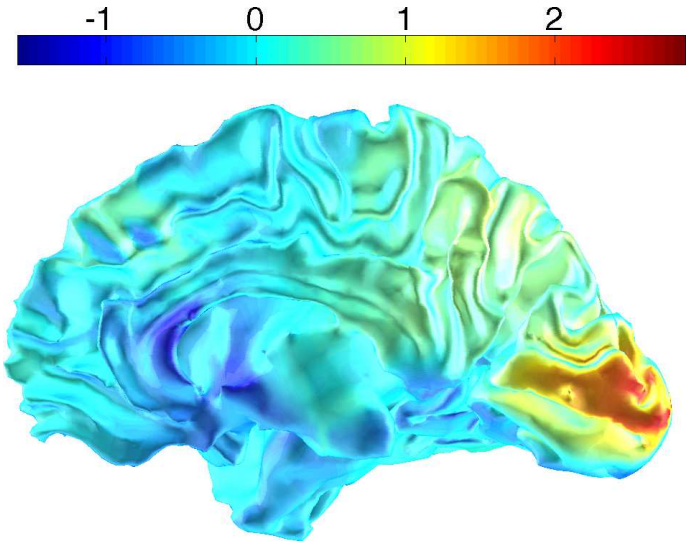


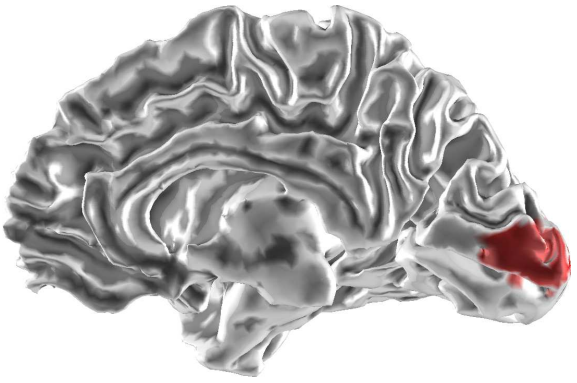
Figure 5.21: Example of T_{exp} map to be thresholded. The color represents at each position i the value T^i_{exp} computed on the Fourier coefficients at $15\ Hz$. Data correspond to the stimulation in the lower left quadrant at a frequency of 7.5 Hz.

5.5.3.2 From localization to retinotopic maps

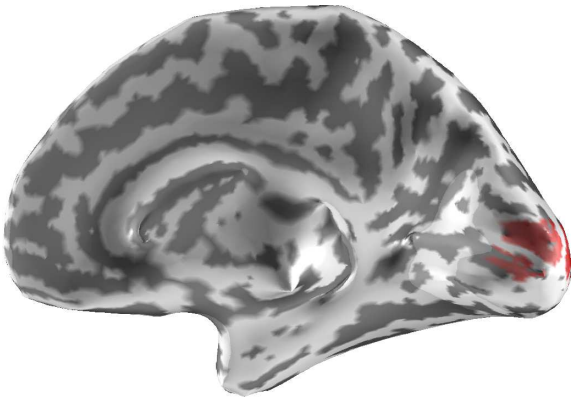
In previous section, we detailed the different intermediate steps to complete in order to obtain a localization result. However, our interest goes beyond simple localization, since our objective is to obtain a pipeline to achieve retinotopic mapping with MEG. This implies that the same parameters should be used for all the experimental conditions, meaning here all the localizations of the flickering pattern in the visual field. We would like to emphasize that this can be relatively challenging and that this issue is rarely mentioned in classical studies where the different experimental conditions are treated separately. This implies for example that the regularization parameter in the inverse problem and that the statistical threshold level should not be manually tuned for each experimental condition. When achieving a mapping like here, all the data for the different conditions are processed in the very same way. Therefore, the processing pipeline needs to be robust to the variations that necessarily occur between different experimental datasets.

In order to address the problem of retinotopic mapping, the results for all the positions in the visual fields need to be displayed on a common source space, *i.e.,* on a same triangulation. When two different conditions both produce a significant activation at a same location, the condition that is selected and displayed is the condition for which the value of T_{exp} is maximum.

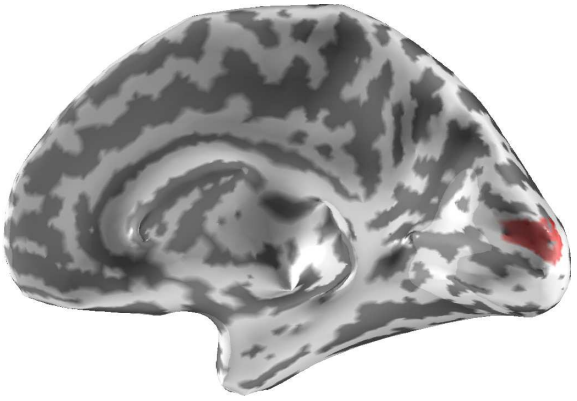
In the following results, the position of the flickering pattern in the visual field is color coded. Color conventions are given in figure 5.23. A result of retinotopic mapping obtained with a minimum-norm is provided in figure 5.24. In order to obtain this result, the regularization parameter in the minimum-norm was set using the 10% rule of thumb from Brainstorm (cf. section 3.2.1.2). The Fourier coefficients were obtained during the steady-state period



(a) Thresholded map on the cortex



(b) Thresholded map on the inflated cortex



(c) Thresholded map on the inflated cortex with only the main connected component lying in V1.

Figure 5.22: Example of thresholded statistical map T_{exp} (p=0.05 with 15000 permutations). Data correspond to the stimulation of an lower left quadrant at a frequency of 7.5 Hz.

in the stimulation time interval (after 400 ms of stimulation) and during the baseline. The P-value was classically set to 0.05 and 15000 permutations were run for each condition.

In this result, it can be observed that the mapping for the left hemi visual field is particularly well recovered on the right hemisphere. The horizontal meridian is correctly mapped in the calcarine sulcus while the upper and lower left visual fields are respectively mapped on the the lower and upper banks. When observing the result for the right hemi visual field, it can be observed that the lower quadrant is correctly mapped on the upper bank of the left calcarine sulcus and that the right horizontal meridian produces an activation that includes the left calcarine sulcus. However, it can also be observed that the Minimum-Norm tends to over estimate the extent of the activation for this particular condition (cf. figure 5.24(a) and figure 5.24(b)).

Please note that this is after observing such results that we investigated the use of the $\ell_{\mathbf{w};212}$ prior to improve the quality of retinotopic mapping with MEG.

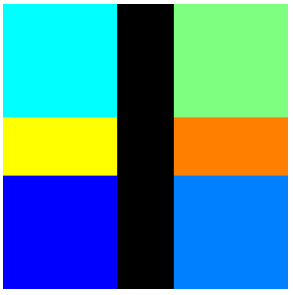


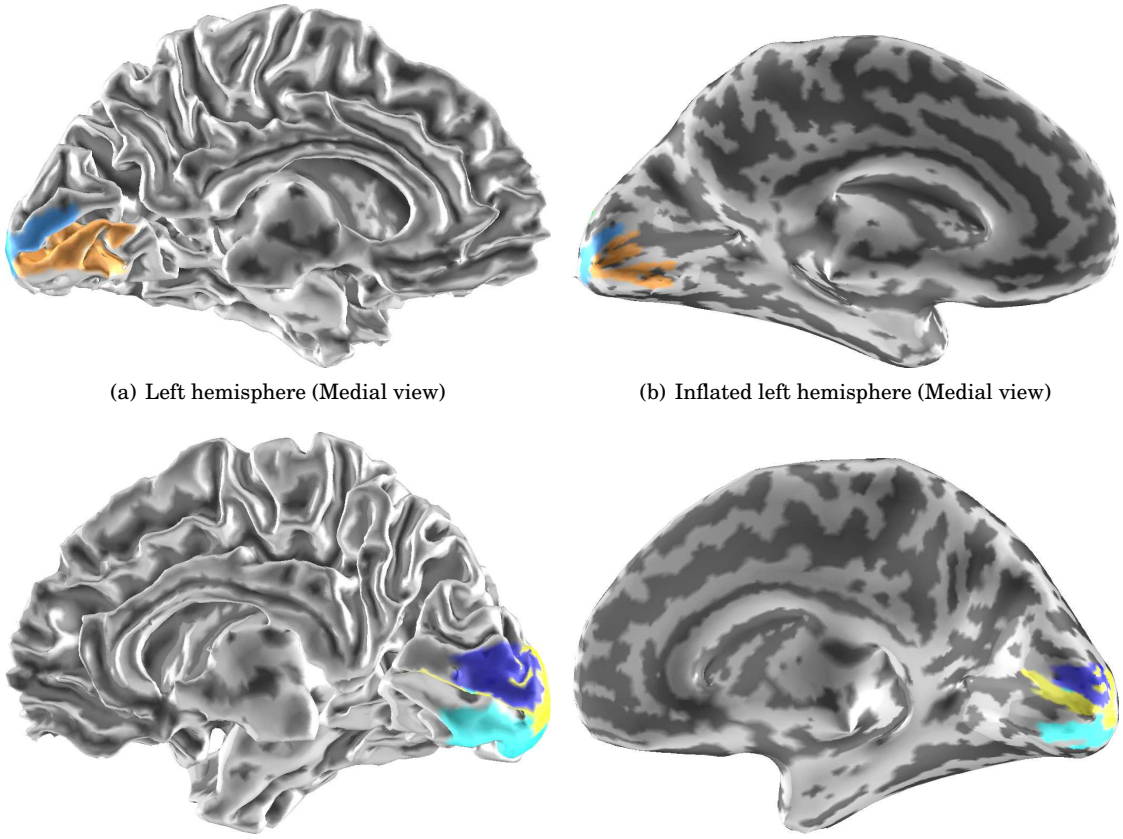
Figure 5.23: Color conventions for each condition represented at their position in the visual field.

5.5.3.3 Reconstruct on WM-GM or GM-CSF interface?

Depending on the M/EEG source analysis pipelines, neural current can be estimated over a mesh separating the gray matter (GM) and the white matter (WM), or over a mesh separating the gray matter from the cerebro-spinal fluid (CSF). This latter interface corresponds to the outer surface of the gray matter. For example the MNE software mentioned at the end of chapter 2 generally presents source estimates on the WM-GM interface while users of the Brainstorm toolbox usually work with the GM-CSF interface extracted with BrainVISA [38]. One reason for this is that the MNE software computes forward models with a 3 layer BEM whose inner layer is the inner skull interface, which is very close to the GM/CSF, while Brainstorm's users work with spherical head models and therefore can use the GM-CSF interface as source space.

During this thesis, we tried our retinotopic mapping pipeline on both interfaces. We present in figure 5.25(a) and figure 5.25(b) two results obtained with the very same parameters for the reconstruction and for the statistical procedure. Some clear differences appear between these two results, which demonstrate the influence of the source mesh on the results. In both cases, the extent of the active region corresponding to the right meridian appears to be overestimated. This is particularly problematic for GM-CSF interface since the active region for the meridian clears the active region for the lower right quadrant. This illustrates particularly well the problem of working with multiple conditions simultaneously when no parameter tuning is performed for each condition individually.

The ℓ_{212} mixed norm presented at the end of chapter 4 is the strategy we investigated during this thesis in order to better control the extent of the active regions.



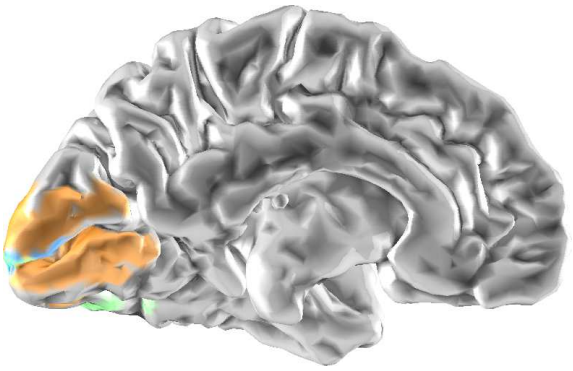
(a) Left hemisphere (Medial view)

(b) Inflated left hemisphere (Medial view)

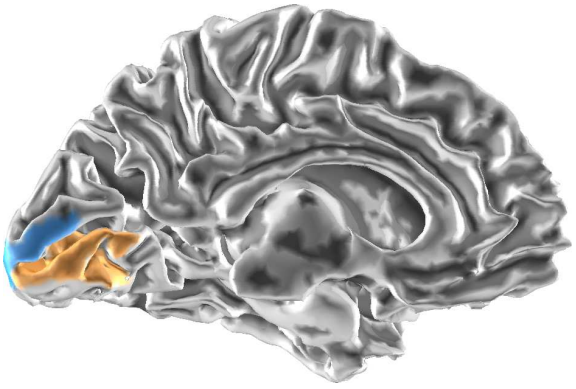
(c) Right hemisphere (Medial view).

(d) Inflated right hemisphere (Medial view).

Figure 5.24: Retinotopic map result obtained using a minimum-norm inverse solver and statistical tests run on the Fourier coefficients (p=0.05 with 15000 permutations). Data correspond to the stimulation at a frequency of 7.5 Hz.



(a) Result on the GM-CSF interface.



(b) Result on the WM-GM interface.

Figure 5.25: Comparison of retinotopic map results obtained by reconstructing with MN on the GM-CSF and on the WM-GM interfaces. Statistical tests were run on the Fourier coefficients (p=0.05 with 15000 permutations). Data correspond to the stimulation at a frequency of 7.5 Hz.

5.5.3.4 Localization results beyond simple ℓ_2 inverse solvers.

The ℓ_2-norm vs. the ℓ_{212}-norm. As it has been observed above, a standard ℓ_2 prior, a.k.a. Minimum-Norm, tends to overestimate the extent of active regions. In order to reduce this problematic behaviour of standard ℓ_2 inverse solvers, we have proposed to invert all the experimental conditions simultaneously and to promote non overlapping activations by using what we called in chapter 4 an inter-condition sparse prior. This prior described in detail in section 4.5 is based on a mixed norm with 3 levels where sparsity is induced between conditions using an ℓ_1 norm. The inverse solver is not linear any more but the problem is still convex, which offers the possibility to perform the current estimation with very efficient algorithms. In the following results, the optimization with the ℓ_{212} prior is performed with Nesterov's iterative scheme (cf. algorithm 4.4 in chapter 4).

When working with an ℓ_{212} prior, the measurements are indexed with a triple index. Here we concatenate the Fourier coefficients estimated on all trials. Each coefficient is indexed by the sensor, the condition and the trial. Compared to what is presented in chapter 4, the last index used here do not correspond to the time but to the trial. With 6 experimental conditions, the matrix to invert is in $\mathbb{R}^{d_m \times 6d_l}$. After running the inverse solver, the estimated Fourier coefficients on the source space form a matrix in $\mathbb{R}^{d_x \times 6d_l}$. The same statistical procedure as for the MN is then used to threshold the activation maps.

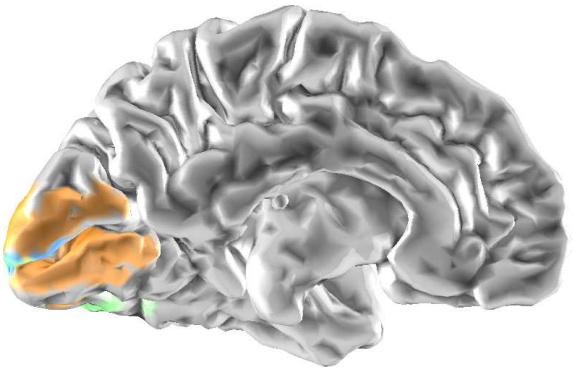
A comparison of mapping results obtained with a simple MN and an ℓ_{212} prior is presented in figure 5.26. It can be observed that the extent of the activation for the right horizontal meridian is significantly improved by using the ℓ_{212} prior. The active region now clearly lies only in the calcarine sulcus which is consistent with our knowledge about the organization of the primary visual cortex. However the activation for the upper right quadrant is still not correctly localized.

The full retinotopic maps obtained with the ℓ_{212} prior are presented in figure 5.27.

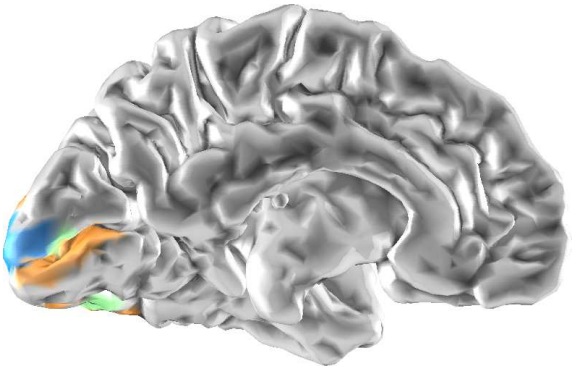
With the MiMS solver. The MiMS inverse solver (cf. section 3.3.1) was also experimented on the same data by Benoit Cottereau during his thesis. Results obtained with his method can be found in [6].

The MiMS solver belongs to the class of Bayesian inverse solvers in the sense that the weights of an ℓ_2 prior are learned. These weights are learned with a multi-resolution approach based on multipolar modeling of spatially extended cortical parcels (cf. section 3.3.1). In the approach exposed in [6], the weights are learned on the full temporal data during the stimulation period. Then the WMN linear inverse solver obtained with the learned weights is used to get the source estimates. Cottereau then estimates the PSD at 15 Hz for each source in each trial using a multitaper approach. The PSD estimator is based on Welch's method (cf. section 5.5.1). Finally, he performs on the PSDs a non-parametric permutation test similar to the one we exposed above.

Our approach and his approach differ in a number of points. First, Cottereau learns on the full data even though he is using only the Fourier coefficient at 15 Hz in his non-parametric statistical test procedure. This means that he may exploit in his localization pipeline information in a wider region of the spectrum than solely 15 Hz. In order to test this hypothesis, it would be interesting to band pass filter the data around 15 Hz to see if the MiMS inverse solver keeps providing the same localization results. Second, by learning the weights with multipolar expansions, Cottereau limits the influence of the dipole orientations on the results. Indeed, the multiresolution approach can be seen as a way to provide regions of interest without really considering the orientation of each individual dipole. Once the matrix used for linear inversion is computed, his strategy is very similar to one we used in this chapter. However, it is necessarily slower since the PSDs are actually estimated in the source space using the reconstructed time series.

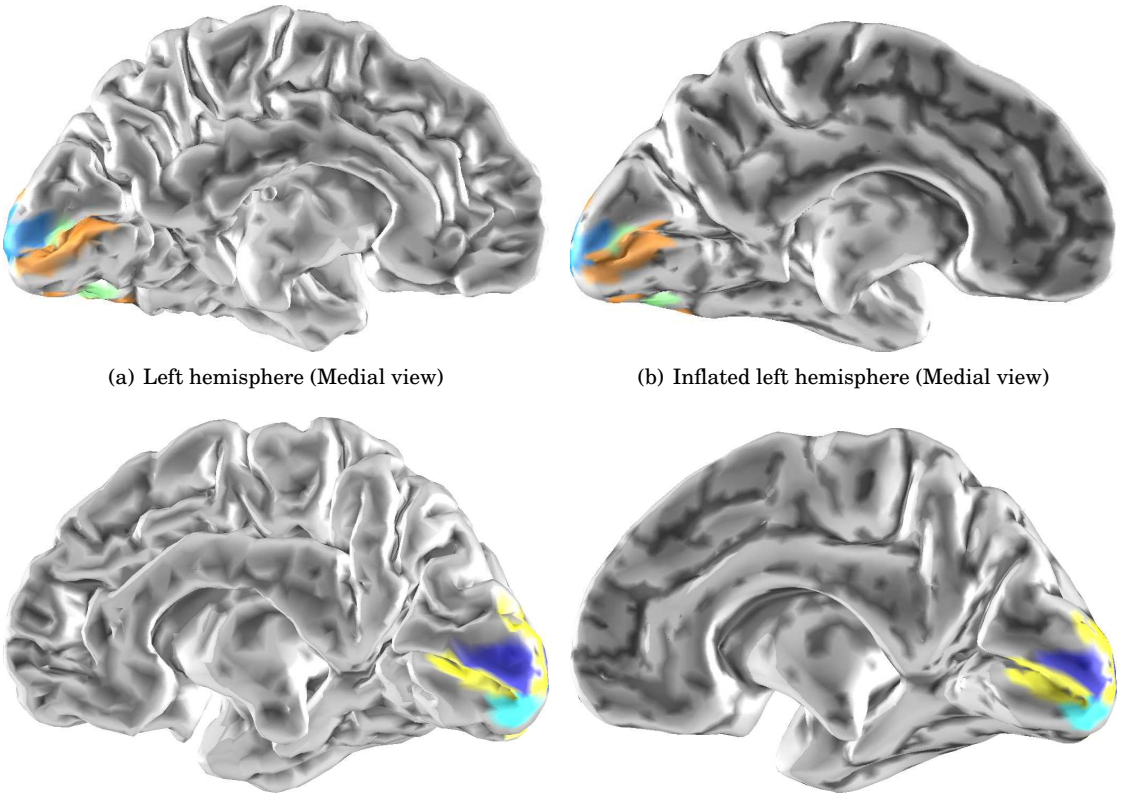


(a) Map obtained with MN.



(b) Map obtained with the ℓ_{212} prior.

Figure 5.26: Comparison of retinotopic mapping results obtained with a MN and with the ℓ_{212} prior. Results are presented on left hemisphere using the GM-CSF interface as source space. Statistical tests were run on the Fourier coefficients (p=0.05 with 15000 permutations). Data correspond to the stimulation at a frequency of 7.5 Hz.



(a) Left hemisphere (Medial view)

(b) Inflated left hemisphere (Medial view)

(c) Right hemisphere (Medial view).

(d) Inflated right hemisphere (Medial view).

Figure 5.27: Retinotopic map result obtained with an inverse solver based on an $\ell_{\mathbf{w};212}$ prior. Statistical tests were run on the Fourier coefficients (p=0.05 with 15000 permutations). Data correspond to the stimulation at a frequency of 7.5 Hz. Results are represented on the interface between the gray matter and the cerebro-spinal fluid.

To our knowledge, Cottereau uses an FFT to extract the PSDs. This is not required if the interest is only in the spectral coefficient at 15 Hz. A simple correlation with a complex sinusoid is enough. It also avoids the problems at the border of the time interval by not requiring to compute a circular convolution like with a discrete Fourier transform.

Taking the best of both approaches would consist in plugging into our pipeline the multiresolution approach in order to improve the MN results with the weights learned by the MiMS procedure. This would favor more spatially regular activation patterns and certainly improve the mapping results while keeping a computationally efficient procedure.

5.5.3.5 Effect of the orientation constraint

As mentioned when discussing the mapping strategy based on MiMS, a possible limitation of the solvers we experienced is their strong dependence on the orientation of the dipoles sampled over the triangulated source space. A solution to circumvent this problem is to work with unconstrained orientations. At each location of the source space lie 3 dipoles, each oriented in the direction of a coordinate axis. This provides 3 coefficients at each location. The amplitude of the activation at each location is then obtained by computing the norm of the vector formed by these 3 coefficients. This provides a straightforward way to estimate the PSD at a given location when no orientation contraints are used. The statistical test procedure can then be identically computed on the PSDs estimated separately on the stimulation and the baseline periods.

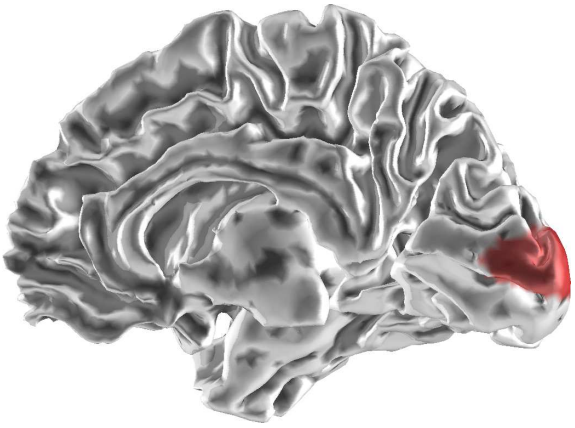
A result obtained with unconstrained orientation on the same data as in section 5.5.3.1 is presented in figure 5.28. It can be observed that the absence of orientation constraints produces spatially smoother active regions. The border of the active region is less influenced by the change of curvature in the source space. Results are consequently more robust to the intricate structure of the cortical region neighboring the calcarine fissure. However, ignoring the orientations tends to produce even wider active regions by spreading the currents estimates over the banks of the calcarine fissure. As a result, retinotopic mapping with an ℓ_2 prior also appeared to be very challenging when using no orientation constraints. Improvement on the robustness of the method to the complex anatomic structure of the occipital cortex has the drawback of an increased tendency of the minimum-norm to smear the reconstructions over wide cortical areas when the orientations are ignored.

5.6 TIMING VISUAL DYNAMICS WITH MEG

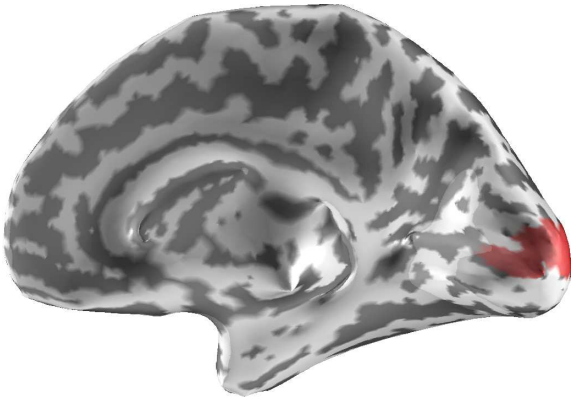
The primary reason for using M/EEG is to exploit its very good temporal resolution. Up to here, we focused on the spatial precision of M/EEG source estimates and we exploited the excellent temporal resolution of M/EEG by using a frequency tagged stimuli. By doing so the SNR is improved which facilitates the localization. However, the ultimate objective concerns the measurement of delays between cortical activations and particularly between different visual areas.

5.6.1 Estimating timings in the visual cortex with M/EEG: Literature review

Using EEG in [213], the authors address this issue with fitted dipoles whose positions were constrained with fMRI localization results. Dipoles are located in V1, jointly V2v and V3v, jointly V2d and V3d, left and right LOV5 (Lateral occipital V5). Once the amplitude time series are estimated for each dipole, the peaks of the waveform can be compared. The delays are estimated by measuring peak to peak time intervals. In [162, 190], delays of activation



(a) Right hemisphere (Medial view).



(b) Right hemisphere (Medial view) on inflated cortex.

Figure 5.28: Example of localization obtained with no orientation constraint using a MN. Statistical tests were run on the PSDs (p=0.05 with 15000 permutations). Data correspond to the stimulation at a frequency of 7.5 Hz in the left lower quadrant of the visual field.

in visual system are also obtained with fMRI localizers and dipole fitting. To our knowledge, timing of activations in the visual cortex with MEG has always been done in conjunction with fMRI data.

In all these studies, the delays observed correspond to a few milliseconds. This observation is of major interest with respect to the frequency used in our steady-state stimulations. Measurements based on the phase are limited to the interval between 0 and 2π, or equivalently $-\pi$ and π. It corresponds to a full cycle which lasts, at a frequency of 15 Hz, 66.6 ms. A measure of phase difference can therefore be related to a time delay as long as the delay of interest is smaller than 66.6 ms. This is the case, in particular, for the delays between activations in the visual cortex as it is reported in various studies [28, 162, 190, 213]. This latter remark suggests that time delays could be extracted from the phase estimated on the sources. A recent study from Di Russo *et al*. [191] tends to confirm this point.

5.6.2 Extracting information from the phase

The Fourier coefficients at 15 Hz at the source level are denoted $\hat{\mathbf{X}}^{stim}$. The coefficient, denoted x_{il}^{stim}, on line i and column l in $\hat{\mathbf{X}}^{stim}$ corresponds to the Fourier coefficient at 15 Hz for the dipole at position i during trial l. In order to evaluate the quality of the phase information, a quantity referred to in the M/EEG literature as *phase lock* can be computed [133].

We define in our case the phase lock, or phase locking value (PLV), for dipole at position i by:

$$\mathrm{PLV}(i) = \frac{1}{d_l}\left|\sum_{l=1}^{d_l}\frac{x_{il}^{stim}}{|x_{il}^{stim}|}\right| \quad .$$

It can be observed that $\mathrm{PLV}(i) \in [0,1]$ and that $\mathrm{PLV}(i)$ is equal to 1 if all the angles, *i.e.,* the arguments, of the complex values x_{il}^{stim} are the same. This means that, the bigger is $\mathrm{PLV}(i)$, the more the phase is stable across trials. This is illustrated in figure 5.29.

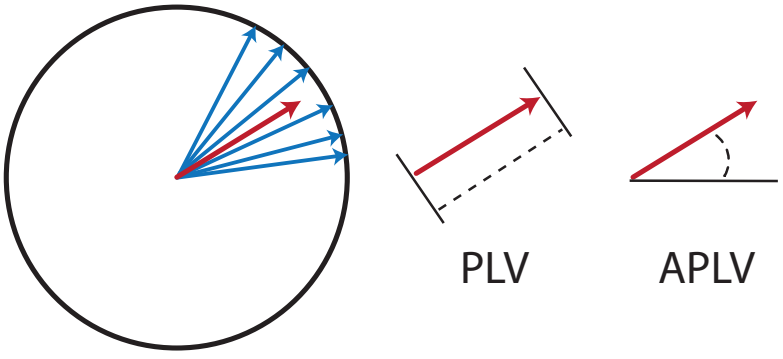


Figure 5.29: Schematic representation to illustrate the computation of the phase locking value (PLV) and the angular information called APLV (see text).

A PLV map computed on the dataset that helped to illustrate section 5.5.3.1 is presented in figure 5.30. A clear “hot spot” can be observed on the upper bank of the calcarine sulcus which is consistent with the mapping results obtained above on the same dataset.

Once the PLV is computed in order to assert that the phase contains information stable across trial, we can investigate the angular part of the average vector. Please note that if a quantity is stable across trials, it means that it is related to the stimulation. While the absolute value of the average vector provides the PLV, the angle contains the phase information

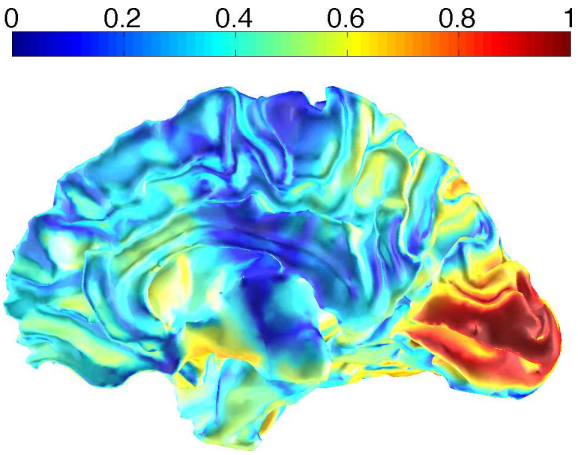


Figure 5.30: Example of phase lock value (PLV) map. The closer is the PLV to 1, the more the phase of the estimated Fourier coefficients is stable across trials. One can observe a clear "hot spot" on the upper bank of the calcarine sulcus. This agrees with our knowledge on V1 for a stimulation in the lower left quadrant in the visual field.

which can provide information about delays. We call this quantity APLV and we define it by:

$$\mathrm{APLV}(i) = \mathrm{ang}\left(\frac{1}{d_l}\sum_{l=1}^{d_l}\frac{x_{il}^{stim}}{|x_{il}^{stim}|}\right) \in [-\pi, \pi] \ .$$

The computation of the APLV is illustrated in figure 5.29.

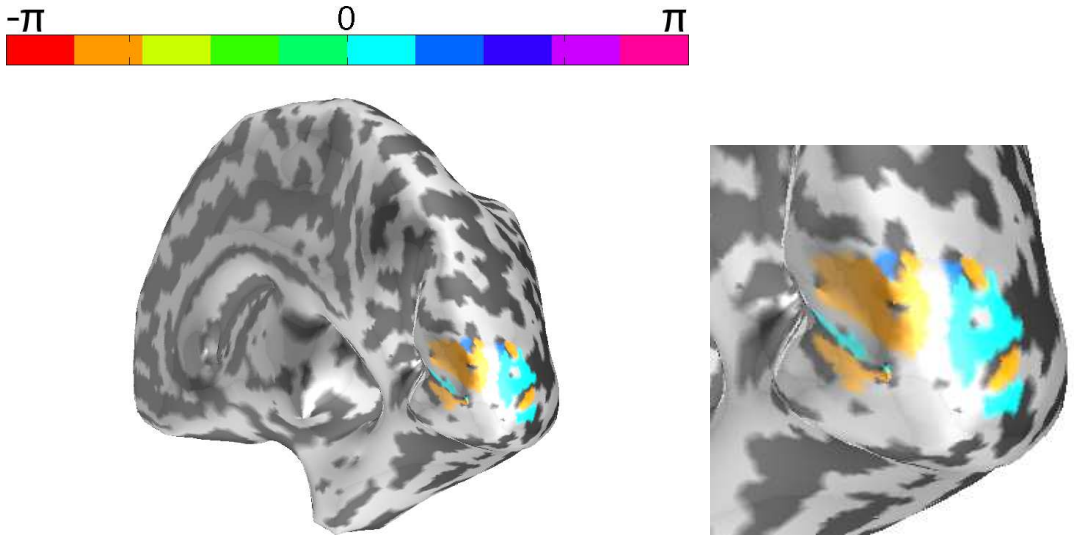
5.6.3 Preliminary results

In previous sections, we detailed our strategy for the analysis of the phase information at the cortical level in order to investigate delays. The PLV provides a principled way to assert if the phase is stable across trials and therefore if it contains information related to the stimulation. If the PLV is close to 1, as it is the case with our example in figure 5.30, the APLV provides an angular information also related to the stimulation which might lead to new insights on the delays. A map of APLV restricted to the regions flagged as active by permutation tests on the Fourier coefficients is provided in figure 5.31.

Such a map shows the presence of distinct values of angle in the occipital region. However, it appears to be hard to directly interpret the differences of angle between two regions as a delay of propagation. Due to the numerous steps to complete in order to obtain such an image, we acknowledge that such an interpretation requires much more validation possibly by testing with multiple stimulation frequencies. Our results on the phase are very preliminary and are presented here just to motivate the study of delay estimation with MEG data and steady-state visual stimulation. Also, we should recall that phase or delay estimation comes after the mapping which means that the mapping procedure should first be considered as a solved problem.

5.7 Discussion

In this chapter, we demonstrated that fast retinotopic mapping of V1 with MEG could be achieved using relatively simple mathematical and algorithmic tools. This was made possible thanks to a set of technical decisions from the design of the protocol to the



(a) Full right hemisphere.

(b) Zoom on occipital cortex.

Figure 5.31: Sample phase map used for delay estimation. The quantity represented is the APLV (see text) restricted to the active region delineated by permutation test on the Fourier coefficients. Data correspond to the stimulation at a frequency of 7.5 Hz in the lower left quadrant of the visual field. Results are represented on the inflated interface between the white matter and the gray matter.

data analysis. The first decision was to use a steady-state stimulation protocol with a flickering pattern. With a steady-state stimulation and a frequency tagged stimulus, it is possible to automatically extract from the spectrum of the signal the relevant information. It is not the case with a standard study based on event related potentials (ERPs). With an ERP based protocol, the information is extracted at the peak of the waveform in the time domain. The peak of activation is not stable across conditions and subjects, therefore the extraction requires a manual intervention. The second advantage of the frequency tagged stimulus is that it allowed us to extract the signal of interest from raw data. All the results presented in this chapter were obtained without any data cleaning. We used for the stimulation a contrasted pattern with multiple orientations in order to increase the amplitude of the neural response and consequently improve the SNR. Finally, we decided to solve the inverse problem with a distributed source model in order to be able to reconstruct the activation pattern produced by an unknown number of active sources but also to obtain mapping results with spatially extended active regions. These concomitant choices of protocol and data analysis strategy allowed us to obtain a fully automatic procedure for the retinotopic mapping of V1 with MEG.

Our understanding of the different steps in the pipeline enabled us to significantly speed up the computation by limiting the computation of the solution of the inverse problem to the Fourier coefficients of interest and by performing the non-parametric tests with simple linear algebra. In this chapter, we argued that to actually achieve retinotopic mapping it is necessary to have a pipeline where all the experimental conditions are processed with the same parameters. In our analysis of the data, we tried to find solutions where manual tuning was not required for each experimental condition independently. This led us to the conclusion that a possible solution was to invert the measurements, or more precisely here the Fourier coefficients, for all the conditions simultaneously.

Such a multi-condition analysis was presented above using a regularization prior based on a $\ell_{\mathbf{w};212}$ mixed norm. In section 5.5.3.4, we demonstrated that this prior can actually improve the mapping results. Our experience with the $\ell_{\mathbf{w};212}$ norm shows that it clearly helps

to delineate the active regions for each condition. It sets the reconstructions to 0 over regions where other conditions are more likely to be active and it enhances the amplitudes of the reconstructions in other regions. By doing so it helps to control the spatial extent of active patterns and it improves the mapping.

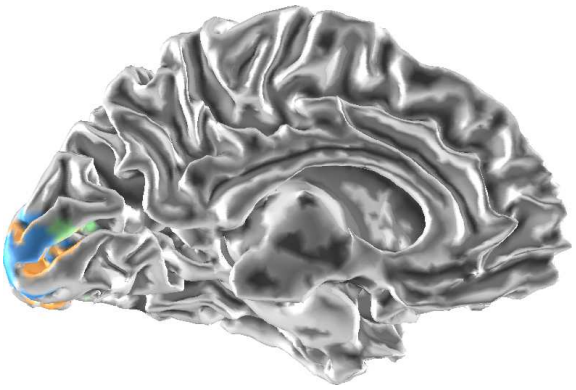
Using the $\ell_{\mathbf{w};212}$ prior implies an increase in computation time. However, the efficient algorithms detailed in chapter 4 allowed us to get results with a $\ell_{\mathbf{w};212}$ prior in a few minutes with highly sampled cortical meshes. However, our experience shows that the $\ell_{\mathbf{w};212}$ prior does not help to obtain significantly active dipoles for conditions where the MN has failed to see something significant.

During our exploration of these data, we met a set of difficulties. Among these is the approximative estimation of the dipole orientations when they are fixed by the normals to the mesh. The cortical region neighboring the calcarine fissure can be very intricate which makes the estimates of the normals to the gray matter quite noisy in this brain region. Sensitivity of the results to the source space is clearly illustrated above, where reconstruction on the WM/GM and GM/CSF interfaces are compared. However, using unconstrained orientations creates other problems already discussed.

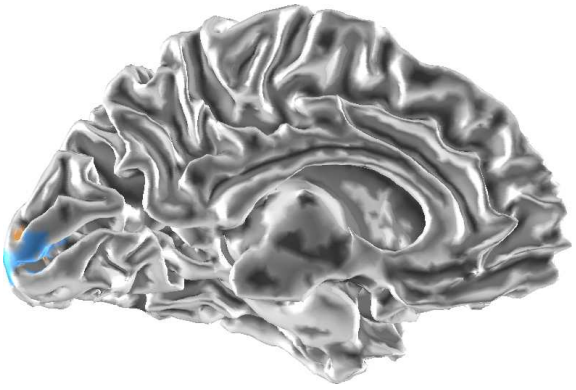
Our interest for other solvers than MN was largely motivated by the difficulties we met when addressing the problem of retinotopic mapping with MEG. In our investigations on these data, we tried to use the ℓ_{21} prior (cf. section 4.4.2) but as we could have expected it did not provide very nice results as this solver by construction sets an ℓ_1 norm over space and consequently cannot really reconstruct spatially extended activations. We also tried to use the Gamma-MAP inverse solver (cf. section 3.3.2) but it appeared quite difficult to have good source covariance templates and a good noise covariance estimate to obtain good results. This is probably due to our limited expertise with this solver on real data. In order to promote spatially smooth reconstructions, we also tried a prior based on the ℓ_2 norm of the surface gradient. We called this solver HEAT in chapter 3. However, results with this solver show that such a prior can significantly change and degrade the localization results when the cortical region of interest is particularly intricate.

In our investigations, we also tried to use other criteria than the 10% rule of thumb from Brainstorm to estimate the lambda in the MN inverse solver. We present in figure 5.32 and figure 5.33 a retinotopic mapping obtained using the GCV and the L-curve methods. The lambda parameter was estimated independently with the Fourier coefficients obtained in each trial (baseline and stimulation). This result presents smaller active regions in comparison to the results in figure 5.24(c). With the GCV, the active region for the upper left quadrant of the visual field (lower bank of the calcarine sulcus on the right hemisphere) almost disappears while with the L-curve it is completely removed. This suggests that the GCV and the L-curve tend both to estimate a regularization parameter smaller than the one obtained with Brainstorm's 10% rule. Also, we can conclude from this example that the L-curve tends to provide a value of lambda smaller than the GCV. This agrees with the recurrent claim that the L-curve approach tends to "under regularize" the inverse problem of M/EEG.

From the investigations and preliminary results presented in this chapter we can draw some conclusions. The first one in that using the phase of the Fourier coefficients when performing statistics can be very useful to improve the mapping results. It actually increases the power of the test. However, our experience tends to prove that estimates of the phase require long periods of stimulation to be estimated robustly. In the results presented above, the phase was estimated on a period of stimulation during 5 s. With the same dataset, we tried our mapping pipeline after artificially shortening the period of stimulation. By doing so, we observed that the mapping quality starts rapidly to degrade when the phase is estimated on less than 3.5 s of MEG signal. We have also run the same computations when working with only the PSDs. It appeared that the mapping is still relatively stable even when the periods of stimulation last 3 s in each trial. This suggests that the phase of the Fourier coefficients

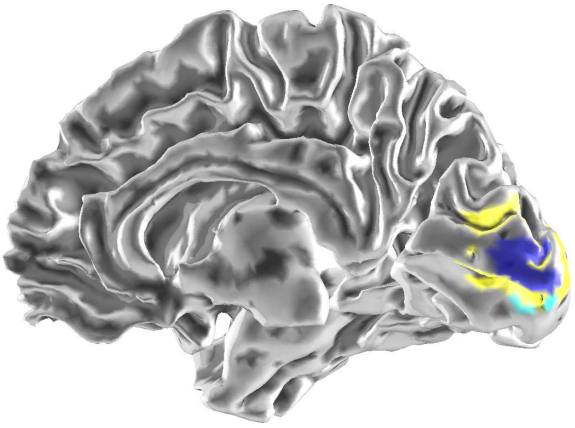


(a) Left hemisphere with GCV.

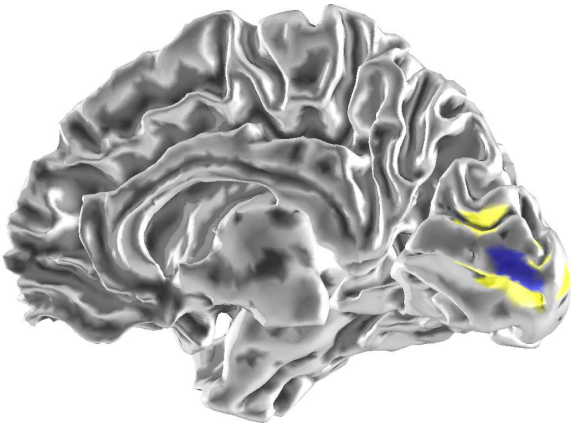


(b) Left hemisphere with L-curve.

Figure 5.32: Comparison of mapping results obtained using the GCV and the L-curve methods to estimate the regularization parameter. Statistics are performed on the Fourier coefficients. Data correspond to the stimulation at a frequency of 7.5 Hz in the right visual field. Results are represented on the WM/GM interface.



(a) Right hemisphere with GCV.



(b) Right hemisphere with L-curve.

Figure 5.33: Comparison of mapping results obtained using the GCV and the L-curve methods to estimate the regularization parameter. Statistics are performed on the Fourier coefficients. Data correspond to the stimulation at a frequency of 7.5 Hz in the left visual field. Results are represented on the WM/GM interface.

can be easily estimated incorrectly.

5.8 Conclusion

While in the literature many studies involve both fMRI and M/EEG data when investigating the human visual cortex, we demonstrated in this chapter that basic retinotopic mapping of V1 was possible with MEG data only. Our review of the literature showed that our approach based on MEG measurements, steady-state stimuli and distributed inverse solvers, had never been done. The steady-state stimuli allowed us to have an automatic way to extract the signal of interest from raw data without any artifact rejection method or filtering. Our expertise on inverse solvers allowed us to design a very efficient pipeline to estimate the Fourier coefficients of interests at the source level. Finally, the non-parametric statistical method we detailed above allowed us to extract significantly active regions very efficiently.

During this study, we faced a set of difficulties. The most important appeared when mapping multiple conditions. What was particularly challenging was the fact that the different conditions *i.e.,* position of the flickering pattern in the visual field, were producing data with differences in quality and SNR. The amplitudes of the signal of interest were different at the sensor level depending on the depth of the source, its orientation and its position (ventral or dorsal). This issue motivated our methodological contribution as a solver where all the experimental conditions are inverted simultaneously. This solver is based on a mixed norm where the overlap of active regions is penalized using an ℓ_1 norm. We called this regularization, an inter-condition sparse prior.

To conclude this chapter, we would like to mention that this study was at the center of our work during this thesis. It motivated most of our methodological investigations and contributions. It was also our first occasion to participate in data acquisitions and the first time we were confronted to the problem of controlling an experimental protocol. We can now fully appreciate the fact that this step of the study is non trivial and particularly important when dealing with brain functional imaging data.

CHAPTER 6

TRACKING CORTICAL ACTIVATIONS WITH SPATIO-TEMPORAL CONSTRAINTS

The work presented in this chapter goes one step beyond inverse modeling and source localization. It aims at providing a way to sketch the evolution of the cortical activations after stimulus onset. The proposed method builds on top of standard and widely used linear inverse solvers and proposes a strategy to extract spatiotemporally consistent, *i.e.,* physiologically relevant, active patterns. The limitations of classical linear inverse solvers are well known. Indeed, as it is discussed in chapter 3, they tend to smear the estimated distributions of currents over the cortex, and they do not achieve a selection of active regions by setting coefficients to zero, as it could be done with a sparsity inducing prior. This work gives a principled method to achieve such a selection and to remove spurious activations that might be introduced by basic instant by instant linear inverse solvers.

Exploiting the graph structure of the triangulated cortical surface and the high time sampling of M/EEG recordings, neural activations are tracked over time using a very efficient graphcut-based algorithm. Such an approach computes a minimum cut on a particularly designed weighted graph, imposing spatiotemporal regularity constraints on the activation patterns. Labels are assigned to each node of the graph, distinguishing between active vs. non-active conditions. The method works globally on the full time period of interest, can cope with spatially extended active regions and allows the active domain to exhibit topology changes over time. The algorithm is illustrated and validated on synthetic data. Application of the method to two MEG datasets demonstrate the ability of the algorithm to track cortical activations in the primary visual cortex and the somatosensory cortex.

Contents

6.1 INTRODUCTION

By providing instantaneous measurements of the weak electromagnetic fields generated by neural activations, M/EEG offer a way to estimate neural currents with a millisecond temporal resolution. Given these current estimates, which can be seen as images of the active brain, a new challenge consists in studying the spatiotemporal evolution of neural activities rather than only localizing specific brain areas involved in experimental tasks [136].

A parallel can be drawn between the challenge proposed here and the problem referred to as tracking in image processing (see [233] for a recent survey). Both applications have some similarities and some differences. With neuroimaging data, activations can rapidly move from one part of the brain to another one via white matter tracks; the signals in the connecting axons are not captured by MEG or EEG. This phenomenon can be compared to the occlusion problem in video sequences. Brain activations can move and appear while their intensities and contrasts change over time, which has some similarity with the *illumination* problem. On the other hand, there are also some fundamental differences. Brain activations are not rigid objects making irrelevant constraints such as point-to-point correspondence between time instants [126], *common motion* constraints or shape priors [138]. The topology and the shapes of the active brain regions can evolve over time. This suggests that the tracking methods developed in the computer vision community can provide principled methods to capture the dynamics of active brain regions, but cannot be applied directly. In the context of brain functional imaging with M/EEG where the activations are defined over a triangulated cortex with a natural graph structure, combinatorial optimization techniques based on graph cuts are the most relevant. Graph cut based techniques in the computer vision community were first applied for image restoration [96] and segmentation [21, 231]. Examples of video segmentation via graph cuts are presented in [21], where a complete set of frames is provided as input to the algorithm that treats the entire sequence as a 3D grid of pixels. Related work such as [122, 232] present results of object tracking using graph cuts by solving the problem frame by frame. In this chapter, a graph cut based approach is designed to achieve the tracking of brain activations and a global optimization on the full temporal data is advocated.

As discussed in chapters 3 and 4, when considering distributed source models, the inverse problem is ill-posed and requires therefore to set priors on the solution. These priors can, for example, be based on the subject's anatomy when constraining the sources to lie on the triangulated cortex, or on results from other imaging modalities such as fMRI. Schematically, setting a prior on the solution consists in defining a norm and finding the solution that has the smallest norm among the ones that explain the measured data sufficiently well.

Commonly, this norm is a ℓ_2 norm. Solutions obtained by such constraints lead to linear solutions (cf. chapter 3). Even if these methods are known to smear the estimated distributions of cortical currents, often leading to solutions that are too widely extended, they are still considered as the standard methods. Technical reasons for the success of such inverse procedures are that they are easy to implement, make very few assumptions on the solutions, are very fast to compute and relatively robust considering the level of noise present in real M/EEG datasets. More importantly, they are used in the M/EEG community because they provide localization results that are sufficiently accurate. There are some cognitive neuroscience studies where the precision of the localization is not critical. On the contrary extreme precision should be required for pre-surgical recordings, *e.g.,* for epileptic patients.

Basic inverse solvers do not integrate constraints on the spatial or temporal regularity of the activations, although physiology imposes such regularity on cortical activations. Various contributions have proposed methods to integrate such smoothness priors in an ℓ_2 inverse problem by adding for example spatial or temporal smoothing operators, like a Laplacian, in the regularization (see section 3.2.2.4). Such Laplacian based methods appear to be diffi-

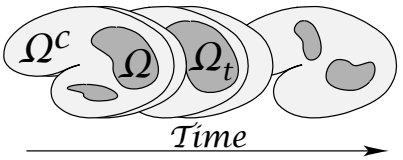
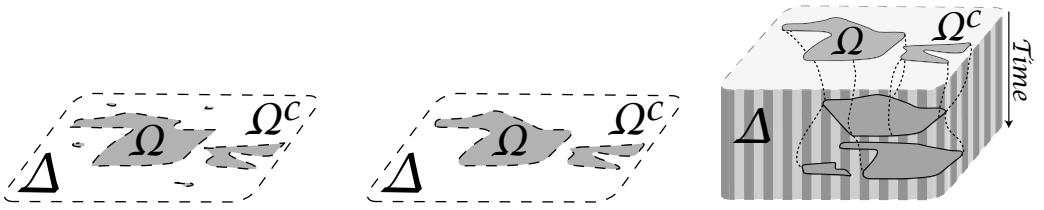


Figure 6.1: Schematic illustration of spatiotemporal active cortical regions. Ω (resp. Ω^c) indicates the active (resp. non-active region). Ω_t is the restriction of Ω to time t.

cult to use with real data as it is unclear how the smoothness prior affects the localization. Depending on the complexity of the cortical sheet structure around the active brain region, a bad spatial smoothness prior can strongly bias the localization result. More recently, the use of a mixed norm has been proposed to better constrain the M/EEG inverse problem (cf. section section 4.4.2). However, such a method, based on an ℓ_1 prior over space and an ℓ_2 prior over time has difficulties to cope with spatially extended activations. With an ℓ_1 prior over space, the lower the SNR, the more the inverse problem is penalized and the more focal is the active region. Also, no matter how good is the SNR, the ℓ_1 prior implies that an optimal solution has fewer active focal sources than the number of sensors. This implies that with a detailed source space an extended region cannot be reconstructed using such a prior. In order to cope with this problem, one could introduce a TV prior with temporal regularization, but such a solver would not be very tractable on real datasets. Since [9], other techniques based on Bayesian estimation have also been applied to the M/EEG inverse problem (cf. section 3.3). Such methods aim at estimating the unknown source covariance matrix in order to learn the good prior used to regularize the inverse problem. Localization precision obtained in simulation studies using these methods is very promising but still these approaches are not flawless. First, these methods rely on the assumption that the source covariance is stable in the time window during which the learning step is achieved. This is not true with real data, especially for long time windows that are considered when looking at late brain responses. Second, the results obtained with these methods depend on the source covariance templates defined as input. Even if the Bayesian estimators developed within these frameworks aim at selecting the good templates, the choice of these templates can have a strong impact on the localization results. We refer the reader to the end of chapter 3 for a discussion on this topic.

Such an observation leads to the conclusion that, even if various recent contributions offer very promising methods to solve the M/EEG inverse problem, standard linear inverse methods, and more specifically simple ℓ_2 priors, provide sufficiently good neural currents estimates, in order to track the dynamics of distributed and rapidly-evolving cortical current patterns in a principled manner. The method presented in this contribution provides a way to follow over time the "hot spots", *i.e.,* the active regions (cf. figure 6.1), while preserving spatiotemporal regularity. In the framework detailed in this chapter, the topology of active regions can change over time. They can appear, split, merge, and disappear. This makes the method able to handle spatially extended active regions while allowing the active domain to evolve during the time window of interest. To our knowledge no other existing method offers such possibilities. Thanks to recent implementations of graph-based algorithms, the method detailed here is tractable on real datasets and offers a very efficient tool to capture the brain dynamics.

The rest of this chapter consists of two parts. Section 6.2 presents the optimization framework that is used to select coherent spatiotemporal activations defined over a triangulated mesh. A variational formulation of the tracking problem is introduced with its discretization over a triangulation, leading to an optimization problem that can be very efficiently solved using a graph cut algorithm. Section 6.2 concludes with a validation on synthetic datasets.



(a) Thresholding (b) Regularized Thresholding (c) Tracking on spatiotemporal domain

Figure 6.2: From thresholding to tracking.

Section 6.3 presents the application of the algorithm to MEG data with two different datasets exhibiting activations in the primary visual cortex and the somatosensory cortex. Even if the results presented are obtained with MEG data, the method can be directly applied to neural currents estimated from EEG data as well.

6.2 Tracking with Graph Cuts on a Triangulated Surface

6.2.1 From Thresholding to Tracking

Let f be a real valued function, defined over a domain Δ

$$f : \Delta \rightarrow \mathbb{R}$$

When Δ contains a temporal dimension, finding an "active" region, denoted Ω, vs. a "non active" region of Δ, denoted Ω^c can be viewed as detecting activity over time. The regions Ω and Ω^c form a partition of Δ, *i.e.,* $\Omega \cap \Omega^c = \emptyset$ and $\Omega \cup \Omega^c = \Delta$. The function f encodes the likelihood for an element of Δ to be inactive, and is thus assumed to take small values in active regions.

A coarse tracking result can be obtained by simple thresholding, *i.e.,* $\Omega^* = \{x \in \Delta \text{ s.t. } f(x) \leq T\}$ where $T \in \mathbb{R}$ is the thresholding value. However, results obtained by thresholding can be very noisy when f is corrupted by noise. Results are considered to be noisy if the border of the active region is irregular or if Ω consists of very small active regions. This is illustrated in figure 6.2(a). It can be shown that the result obtained by thresholding is the solution of the following variational problem:

$$\Omega^* = \arg\min_{\Omega} \int_{\Omega} f(x)dx + \int_{\Omega^c} T dx = \arg\min_{\Omega} \mathcal{D}(\Omega) \ . \tag{6.1}$$

$\mathcal{D}(\Omega)$ is a *data fidelity* term. One can improve the thresholding by forcing the solution to be *regular*. This is done using a Lagrangian approach that adds a term to (6.1) that penalizes solutions Ω based on a measure of their regularity $\mathcal{R}(\Omega) \in \mathbb{R}_+$. Equation (6.1) becomes

$$\Omega^* = \arg\min_{\Omega} \mathcal{D}(\Omega) + \lambda\mathcal{R}(\Omega) \ , \ \lambda \in \mathbb{R}_+ \ . \tag{6.2}$$

To improve robustness to noise, the regularity measure should prevent the occurrence of small isolated regions. If the domain Δ is only spatial, such a regularity can be enforced by

penalizing the solution Ω^* by the length of its border $\partial\Omega^*$ [158]. Figure 6.2(b) illustrates a result obtained with such a regularization.

If Δ is a spatiotemporal domain, imposing the regularity of $\partial\Omega$ can be achieved by enforcing the restriction of the domain at each time instant to have a small perimeter but also by enforcing an overlap between neighboring time restrictions of Ω. The regularisation measure $\mathcal{R}(\Omega)$ can be separated in two parts: a spatial regularisation measure, denoted $\mathcal{R}_{space}(\Omega)$, and a temporal measure, denoted $\mathcal{R}_{time}(\Omega)$. A solution obtained using the penalty term $\mathcal{R}(\Omega) = \mathcal{R}_{space}(\Omega) + \mathcal{R}_{time}(\Omega)$ is illustrated in figure 6.2(c). By imposing to the active region, Ω^*, to be regular over space and time, one creates the *tubular structures* that appear in figure 6.2(c). Each of the tubular structures can be seen as an active region evolving over time. Being able to exhibit such tubular structures on a triangulated domain, and therefore to *track* the activations over time, is the objective of the algorithm proposed in this chapter.

6.2.2 Discretization on a Triangulation

Let us consider $\mathbb{T}$ a triangulation consisting of vertices x_i and triangles n_p, and f a function defined over the vertices of $\mathbb{T}$ and over time:

$$f : (x_i)_i \times (t_k)_{k=1,\dots,K} \to \mathbb{R} \tag{6.3}$$

The set of pairs of adjacent triangles is denoted by $\mathcal{E}$, *i.e.,* "n_p and n_q are adjacent triangles" is equivalent to "$(p,q) \in \mathcal{E}$ and $(q,p) \in \mathcal{E}$". The restriction of f to an instant t_k is denoted by f_{t_k}. To clarify the presentation, K is first supposed to be equal to 1, $f = f_{t_1}$. It corresponds to the case where Δ is only spatial, *i.e.,* no temporal dimension. On a triangulation, partitioning Δ in Ω^c and Ω consists in assigning to each triangle n_p a label 0 or 1. The label 0 corresponds to Ω^c and 1 to Ω. The integrals in (6.2) can be rewritten:

$$\int_\Omega f(x)dx = \int_\Omega f_{t_1}(x)dx = \sum_{n_p \in \Omega} \int_{n_p} f_{t_1}(x)dx \text{ and } \int_{\Omega^c} T dx = \sum_{n_q \in \Omega^c} \int_{n_q} T dx$$

The perimeter of the active region is obtained by the discretization of $\partial\Omega$ using the edges of the triangulation. The regularization term is here given by the sum of the lengths of the edges separating Ω from Ω^c:

$$\int_{\partial\Omega} dl = \sum_{(p,q)\in\mathcal{E}/n_p\in\Omega,\ n_q\in\Omega^c} l_{pq}$$

where l_{pq} stands for the length of the edge between triangle n_p and triangle n_q. Furthermore, this regularization term is weighted by a constant defined as λ_{space}. The energy in (6.2) becomes:

$$\Omega^* = \arg\min_\Omega \sum_{n_p\in\Omega} D_p(1) + \sum_{n_q\in\Omega^c} D_q(0) + \lambda_{\text{space}} \sum_{(p,q)\in\mathcal{E}/n_p\in\Omega,\ n_q\in\Omega^c} l_{pq} \tag{6.4}$$

where $\int_{n_p} f_{t_1}(x)dx = D_p(1)$, and $\int_{n_q} T dx = D_q(0)$ (0 and 1 refer to the 2 labels). If f is assumed to be affine on each triangle, i.e, f is discretized with P1 elements, $D_p(1) = a_p(f(x_{p_1}) + f(x_{p_2}) + f(x_{p_3}))/3$ where p_1, p_2 and p_3 are the indices of the vertices of the triangle n_p, and a_p stands for its area. Similarly, $D_q(0) = a_q T$.

By rewriting (6.2) in this discrete form, the energy to minimize has been cast into a Markov Random Field optimization framework [84] that can be very efficiently solved using graph-based methods [179]. These methods, that have recently been extensively used in Computer Vision [23], establish the equivalence between energy minimization and finding

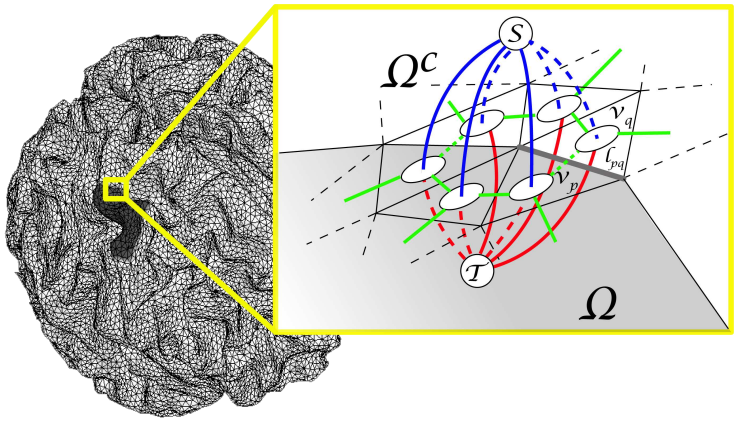


Figure 6.3: Energy discretization on a triangulated mesh.

Table 6.1: Edge weights, *i.e.,* link capacities, of the graph for tracking on a triangulated mesh. Graph nodes n_p are indexed by a space index p. N-Links of type "Spatial" control spatial regularization.

T-Links	Weight
$\mathcal{S} \rightarrow n_p$	$D_p(0)$
$n_p \rightarrow \mathcal{T}$	$D_p(1)$

N-Links	Weight	Type
$n_p \leftrightarrow n_q$	$\lambda_{\text{space}} l_{pq}$	Spatial

the *minimal cut* of a specially designed graph. They are commonly known as "graph-cut" methods. The difficulty is to design a weighted graph providing a natural correspondence between the partitioning of the graph and the energy that is minimized.

See Appendix B for a short introduction to graph cuts.

An illustration of the graph constructed for the current optimization is presented in figure 6.3. Such a construction is inspired by [96], where a similar graph is used for binary image restoration. Contrarily to [96], or more recently to [21], where the graph is constructed on nD grids, the current application imposes to work on temporal data defined on 2D triangulated surfaces embedded in 3D.

Each triangle of the cortex mesh corresponds to one node of the graph. These nodes are thus indexed like the triangles, *i.e.,* using the notation n_p. There are two supplementary terminal nodes : the "Source" $\mathcal{S}$ and "Sink" $\mathcal{T}$ that represent respectively the domains Ω and Ω^c. Each node n_p is linked to both $\mathcal{S}$ and $\mathcal{T}$ (these edges referred as T-links imply that the triangle n_p can belong to one of the two domains represented by $\mathcal{S}$ or $\mathcal{T}$). Furthermore, the graph contains edges between each of the nodes that correspond to adjacent triangles. These edges are referred as N-links. Cutting the graph in two consists in separating the "Source" from the "Sink" by removing some edges. With a minimal cut, each node will remain connected to only one of the terminal nodes so that the remaining graph directly corresponds to a partitioning of the mesh into two domains Ω and Ω^c. Table 6.1 details the edge weights corresponding to the energy (6.4). In practice, edge weights, *i.e.,* link capacities, must be positive. Therefore, prior to the computation of the minimum cut, edge weights are translated to guarantee that they all satisfy this computational constraint.

One can notice that the cost associated to a cut, defined as the sum of the edge weights along the path of the cut, is equal to an energy value. The minimum cut thus provides the optimum. Graph partitioning via minimum cut is, in turn, known to be equivalent to a polynomial problem: the *max flow* problem [73]. The fact that an exact solution is obtained by a single binary cut guarantees the algorithm to be extremely fast (a few seconds for the problems addressed in this contribution) and also globally optimal [179]. In practice, the *min-cut*

Table 6.2: Edge weights, *i.e.,* link capacities, of the graph for tracking on a triangulated mesh. Graph nodes $n_{p,k}$ are indexed by a space index p, and a time index k. "Spatial" N-Links control spatial regularization and "Temporal" ones control temporal overlap between neighboring time instants thus temporal regularization. a_p is the area of triangle p and l_{pq} is the length of the edge separating triangle p and triangle q.

T-Links	Weight
$\mathcal{S} \rightarrow n_{p,k}$	$D_{p,k}(0)$
$n_{p,k} \rightarrow \mathcal{T}$	$D_{p,k}(1)$

N-Links	Weight	Type
$n_{p,k} \leftrightarrow n_{q,k}$	$\lambda_{\text{space}} l_{pq}$	Spatial
$n_{p,k} \leftrightarrow n_{p,k+1}$	$\lambda_{\text{time}} a_p$	Temporal

is therefore obtained via the computation of the *max-flow* using an open source implementation[1] [22].

When considering multiple time instants, a similar approach is used. The nodes $(n_{p,k})_{p,k}$ are now indexed by the triangle index p and the time k. The full graph is obtained by stacking spatial graphs obtained for each t_k and adding N-links between triangles in neighboring time instants. The number of nodes in the graph is now equal to the number of time instants times the number of triangles. Both terminal nodes $\mathcal{S}$ and $\mathcal{T}$ are still unique. Edges weights can now integrate temporal smoothness (See Table 6.2).

In the optimization framework detailed above, the smoothness term becomes $\lambda_{\text{space}}\mathcal{R}_{space}(\Omega) + \lambda_{\text{time}}\mathcal{R}_{time}(\Omega)$, where:

$$\begin{aligned} \mathcal{R}_{space}(\Omega) &= \sum_{k=1}^{K} \mathcal{R}(\Omega_{t_k}) = \sum_{k=1}^{K} \left(\sum_{(p,q)\in\mathcal{E}/n_{p,k}\in\Omega_{t_k},\ n_{q,k}\in\Omega^c_{t_k}} l_{pq} \right) \\ \mathcal{R}_{time}(\Omega) &= \sum_{k=1}^{K-1} \left(\sum_{p/n_{p,k}\in\Omega_{t_k}, n_{p,k+1}\in\Omega^c_{t_{k+1}}} a_p + \sum_{p/n_{p,k}\in\Omega^c_{t_k}, n_{p,k+1}\in\Omega_{t_{k+1}}} a_p \right) . \end{aligned} \tag{6.5}$$

In this formulation the regularization parameters λ_{space} and λ_{time} do not depend on the position in space or in time. However they could be tuned independently for each edge (p, q). One could think of using for example the curvature of the cortex to promote the cuts on regions of high curvature. However, without such a priori and for the sake of simplicity λ_{space} and λ_{time} were kept constant over space and over time.

The weights are detailed in Table 6.2. N-Links of type "Temporal" promote overlap between neighboring time instants and thus enforce temporal smoothness.

The complexity of the graph cut algorithm is $O(N^3)$ where N stands for the number of nodes in the graph. However in practice, like observed in [22] with nD grids, computation time appears to increase linearly with the number of nodes (cf. figure 6.4). More than the computation time, the limiting factor when dealing with large graphs is the memory consumption of the implementation used in this contribution.

6.2.3 Tracking Results with Synthetic Data

The tracking algorithm is now illustrated on two synthetic datasets. The first simulation on a randomly triangulated sphere is designed to be simple and to demonstrate the influence of the regularization parameters. The algorithm is then applied to a more realistic dataset exhibiting three simultaneous moving "hot spots" on a *Bunny* triangulation.

For the first dataset, the domain Δ consists of a triangulated sphere with 3 time instants. About 30 000 vertices are randomly sampled over the sphere and the triangles are obtained with a Delaunay triangulation. The f function was generated to simulate the displacement

[1] http://www.adastral.ucl.ac.uk/~vladkolm/software.html

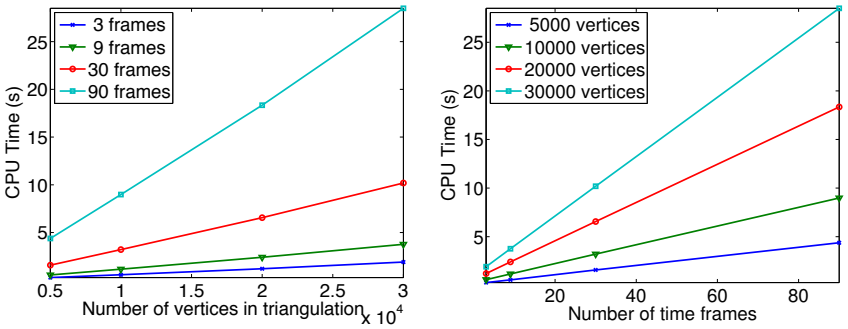


Figure 6.4: Computation times measured on a synthetic dataset. The computation time of the tracking algorithm appears in practice linear with the number of vertices in the mesh (a) and the number of time frames (b). Computation was run on an Intel Core 2 Duo 2.3 GHz CPU with 2 GB of RAM.

of an activation over time, with the addition of a small active region only at the second time instant (see figure 6.5(a)). The function f taking the values 0 in active regions and 1 outside was then corrupted by an additive Gaussian noise with a standard deviation equal to 1.

The tracking algorithm was applied to the data with a threshold T equal to 0.5. Results are presented in two conditions: first, in figure 6.5(c), with only the spatial regularization constraint, *i.e.,* $\lambda_{\text{space}} = 2$ and $\lambda_{\text{time}} = 0$, and second, in figure 6.5(d), with both spatial and temporal constraints active, $\lambda_{\text{space}} = 2$ and $\lambda_{\text{time}} = 0.1$. It can be observed that $\lambda_{\text{space}} > 0$ induces spatially coherent regions while $\lambda_{\text{time}} > 0$ causes the small region only present in frame 2 to disappear. It can also be noticed that the result in figure 6.5(b) obtained with simple thresholding is extremely noisy. The method actually manages to select spatiotemporally consistent activations.

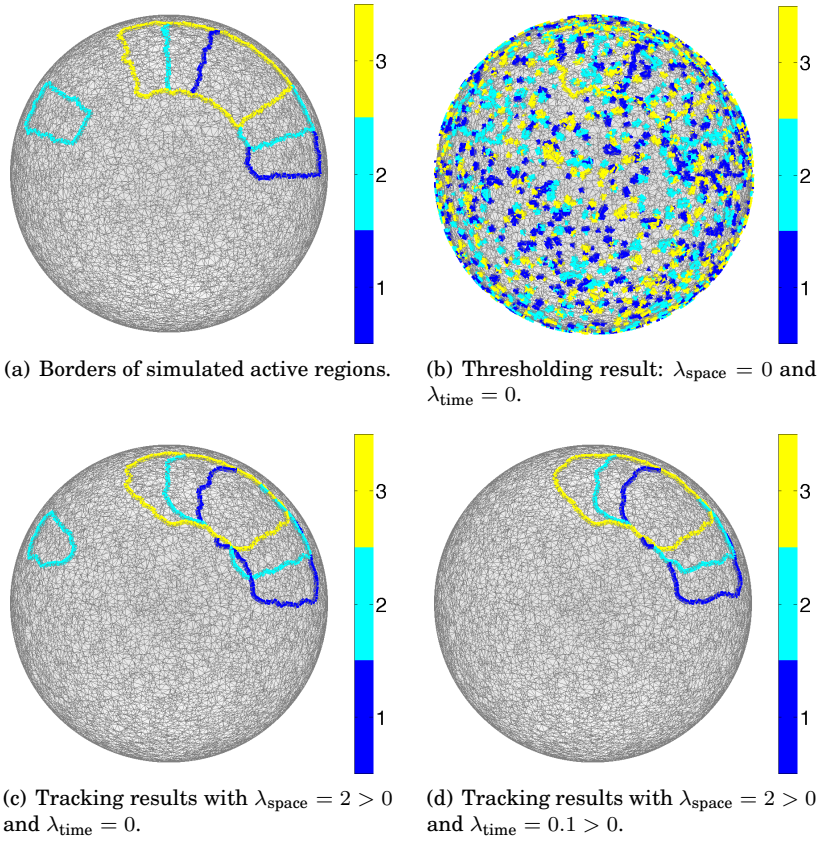
In order to evaluate the sensitivity of the method to the choice of the regularization parameters, a simulation study has been performed. The computation has been run multiple times with various pairs of parameters $(\lambda_{\text{space}}, \lambda_{\text{time}})$. For each pair, the result was compared to the ground truth that was used to simulate the data, i.e, the active region in figure 6.5(a) without the small false positive region in frame 2.

The error was quantified with 3 different measures. The first one is given by the ratio between the number of mislabeled vertices and the total number of vertices, here $30000 * 3$ (cf. figure 6.6). It can be observed with figure 6.6(b) that the method provides accurate results with parameters in a wide range around the optimal obtained with $\lambda^*_{\text{space}} = 2.3$ and $\lambda^*_{\text{time}} = 0.04$. The second performance measure is based on the number of connected components obtained in the result. The right number of connected components is 3, one per time frame. The number of connected components for each pair of parameters is provided in figure 6.6(c). One can observe that the right number of components is correctly estimated for a large range of parameters. Finally, error in active areas is quantified using the Dice's coefficient (DC) between two domains Ω and Ω':

$$DC(\Omega, \Omega') = 2\frac{\text{area}(\Omega \cap \Omega')}{\text{area}(\Omega) + \text{area}(\Omega')} \tag{6.6}$$

which ranges from 0 (no overlap) to 1 (perfect overlap). One can observe in figure 6.6(d) that the Dice's coefficient stays close to 1 for a large range of parameters around $(\lambda^*_{\text{space}}, \lambda^*_{\text{time}})$.

These observations confirm that the method is robust to an approximate definition of the parameters. This is also confirmed by the following results on synthetic and real MEG data,



(a) Borders of simulated active regions.

(b) Thresholding result: $\lambda_{\text{space}} = 0$ and $\lambda_{\text{time}} = 0$.

(c) Tracking results with $\lambda_{\text{space}} = 2 > 0$ and $\lambda_{\text{time}} = 0$.

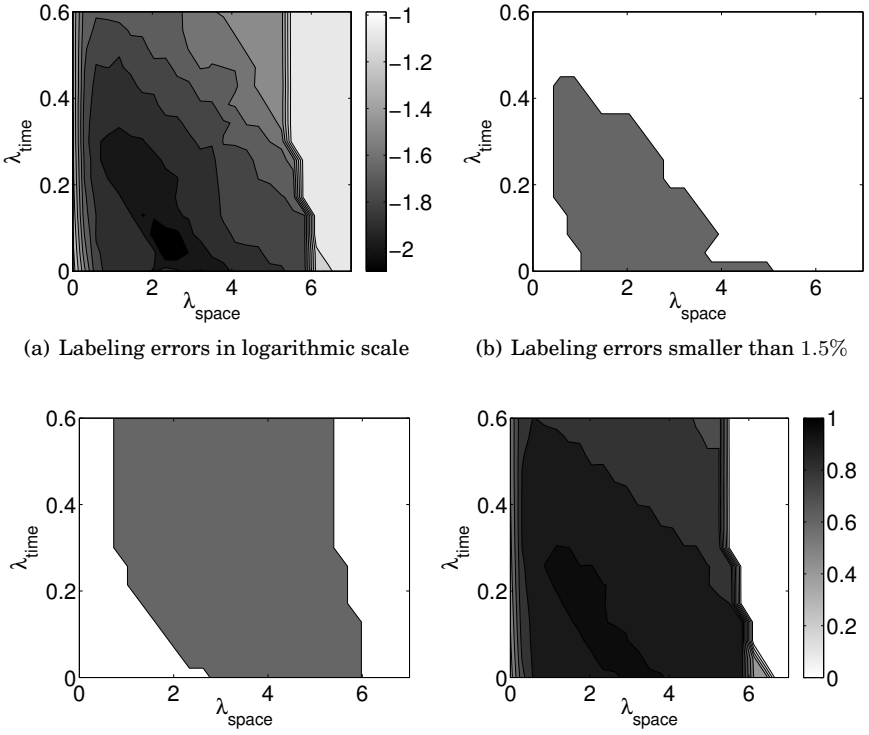
(d) Tracking results with $\lambda_{\text{space}} = 2 > 0$ and $\lambda_{\text{time}} = 0.1 > 0$.

Figure 6.5: Result of tracking using the graph cut algorithm on synthetic dataset defined on a randomly triangulated sphere with 30 000 vertices. Colored lines correspond to the border of the active regions and the color codes for the time instant. Initial data f represented in (a) is equal to 0 in active regions and 1 outside. Prior to the tracking, Gaussian white noise with a standard deviation equal to 1 was added to f.

for which the definition of regularization parameters never actually required a very fine tuning.

In a second dataset, three "hot spots" are moving simultaneously during 100 time frames over a "Bunny" triangulation with about 8000 vertices. These data, also used as a validation set in [135], are presented at 5 different time instants in figure 6.7(a). Such data were designed to provide a more complex and realistic synthetic dataset according to the geometry but also to the time scales of the brain activations measured by M/EEG. In order to respect the convention that f should take small values in active region, f was set to the opposite of the actual activations defined over the mesh. According to the signal amplitude, the parameter T was set to $-4 * 10^{-3}$. Prior to the tracking, an additive Gaussian white noise with standard deviation equal to $6 * 10^{-3}$ was added to the synthetic data. Results of thresholding and tracking are presented in figure 6.7(b) and figure 6.7(c). It can be noticed that the method can actually cope with topology changes since figure 6.7(c) presents the merging of two active regions. Here also, it can be observed that the tracking algorithm provides a clear view of the dynamics of the activation defined over the triangulation.

6.3 Application to M/EEG Data



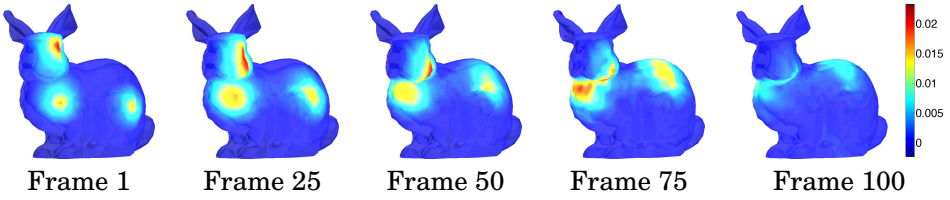
(a) Labeling errors in logarithmic scale

(b) Labeling errors smaller than 1.5%

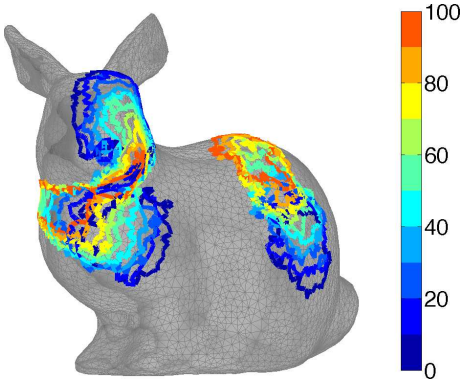
(c) Region where the estimated number of components is equal to 3

(d) Dice's coefficient

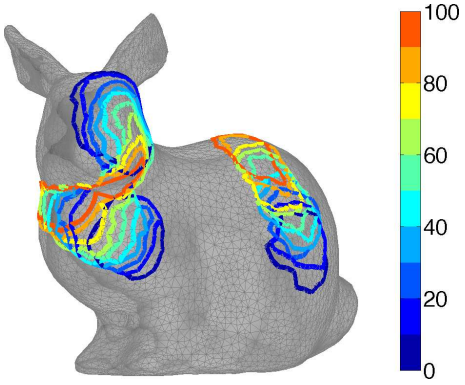
Figure 6.6: Labeling errors obtained by the tracking algorithm for various pairs of regularization parameters $(\lambda_{\text{space}}, \lambda_{\text{time}})$. In (a) Error was quantified by the ratio between the number of mislabeled vertices and the total number of vertices. The color-coded errors are presented in logarithmic scale. The best performance is obtained with $\lambda_{\text{space}} = 2.3$ and $\lambda_{\text{time}} = 0.04$, but the performances remain very acceptable with parameters in a wide interval around these values. This is illustrated in (b) where is represented the region in which errors are smaller than 1.5%. In (c) is represented the region where the number of components is correctly estimated to 3. In (d) the performance is measured using the Dice's coefficient (6.6). The closer it is to 1, the better it is. All performance measures confirm that the result is relatively robust to the definition of the parameters λ_{space} and λ_{time}.



(a) Synthetic dataset on "Bunny" triangulation presented at multiple time instants.



(b) Tracking result with no regularization : $\lambda_{\text{space}} = 0$ and $\lambda_{\text{time}} = 0$



(c) Tracking result with spatiotemporal regularization : $\lambda_{\text{space}} > 0$ and $\lambda_{\text{time}} > 0$

Figure 6.7: Result of tracking using the graph cut algorithm on a synthetic dataset defined on the "Bunny" triangulation with 8000 vertices and 100 time instants. The original data consists of three moving "hot spots" illustrated in (a). Data were corrupted by an additive Gaussian white noise with a standard deviation equal to $6 * 10^{-3}$. Figure (c) demonstrates the ability of the method to cope with topology changes. Between frame 0 and approximately frame 30 the 2 activations on the head of the bunny merge.

The tracking method presented above is now applied on two MEG datasets. The first one is obtained with a visual stimulation paradigm that consists of a series of expanding checkerboard rings. Such a stimulation creates a propagation of activation along the primary visual cortex (V1) that enables direct application of the tracking algorithm. The second dataset consists of a somatosensory finger stimulation. For this dataset, due to the different amplitude levels of activations within the various brain regions involved in the processing of the task, a particular data fidelity cost is designed prior to the tracking algorithm.

6.3.1 Results on visual stimulation

In this experiment, expanding checkerboard rings extended radially from 0 to 4 degrees of visual eccentricity are presented periodically with a frequency of 5 Hz (see figure 6.8). Because of the retinotopic organization along the calcarine fissure and V1, the optical flow generated by the expanding checkerboard rings produces a posterior - anterior wave of activation as illustrated by figure 6.9. It is this propagating wave within V1 that we propose to track.

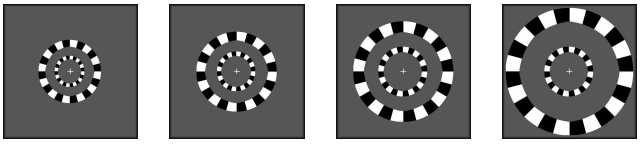


Figure 6.8: One block of successive frames used to produce expanding checkerboard rings. The block of 4 frames is projected periodically with a frequency of 5 Hz.

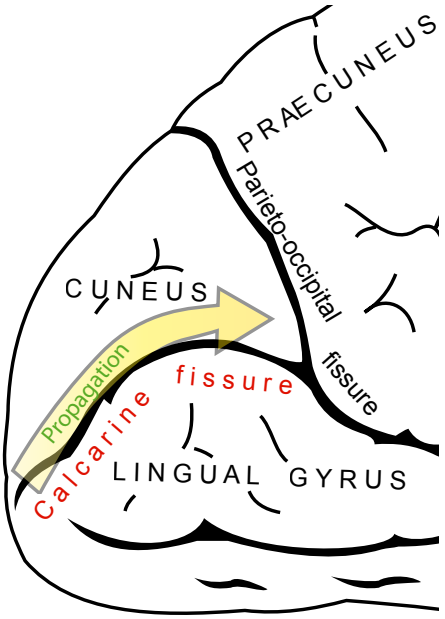


Figure 6.9: Schematic representation of the cortical activation propagation produced by the expanding checkerboard rings. The propagating wave covers the primary visual cortex (V1) on both sides (superior and inferior) of the calcarine fissure, from the posterior to the anterior part of the cortex (Adapted from 20th U.S. edition of Gray's Anatomy of the Human Body, 1918, public domain).

Data acquisition and analysis

The MEG data were acquired at 1250 samples/sec with a 151-SQUID sensor CTF MEG (CTF System Inc.). In each trial, rings were presented for 2.5 s after a prestimulation period of 1.4 s. The first 500 ms of stimulation correspond to a transitory period before 2 s of steady state period (cf. figure 6.10). In order to improve the signal-to-noise ratio, 33 trials were averaged and the data were band pass filtered between 2.5 and 7.5 Hz. These data come originally from the study presented by D. Cosmelli *et al*. in [41].

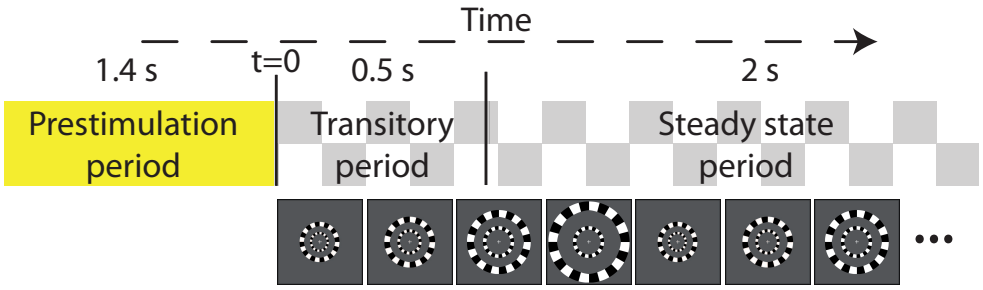
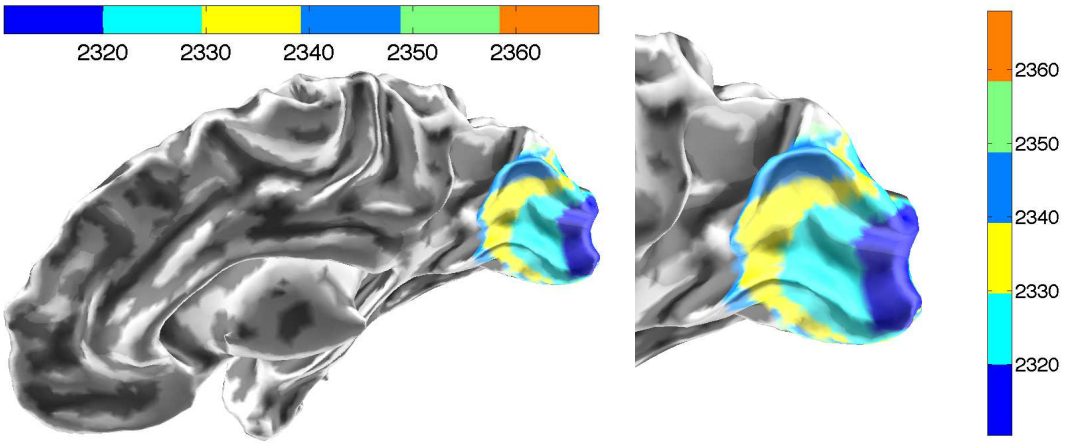


Figure 6.10: Experimental protocol for visual stimulation with the expanding checkerboard rings. Prestimulation period lasts for 1.4 s before 2.5 s of stimulation. Stimulation period is divided in two: the transitory period estimated around 0.5 s and the steady state period that lasts for 2 s. Tracking is performed during the steady state period.

In order to estimate the source amplitudes, the forward problem was computed with a spherical head model and a distributed source space consisting of a cortex triangulation with about 50 000 vertices. An inverse solution was computed with an ℓ_2 penalization term using dipoles with unconstrained orientations. The source amplitudes, denoted s_{it} (i indexes space and t time), were computed by computing the norm of the equivalent current dipole at each location. Using prestimulation recordings s_{it} was then normalized. Assuming stimulation starts at $t = 0$, this corresponds to: $z_{it} = s_{it}/\sigma_i$ where σ_i is the standard deviation estimated on $(s_{it})_{t<0}$. By doing so, the z_{it} are all positive and have large values in active brain regions.

Tracking

The tracking algorithm was run using as data term the computed z_{it}. To follow previously exposed conventions it leads to $f_t(i) = -z_{it}$. The threshold T was set manually in order to obtain active regions lying approximately within V1. To limit computation time and memory consumption during processing, the tracking was performed independently on each period of stimulation during the steady state period. With the 5 Hz stimulation frequency, it corresponds to a period of 200 ms. The results of the tracking algorithm are presented in figure 6.11 while a comparison between thresholding and tracking results is presented in figure 6.12. It can be observed in figure 6.11 a clear propagation of the activation along the calcarine fissure and V1, from the posterior part of the cortex towards its anterior part. By observing the detailed comparison in figure 6.12, it appears clearly that the graph cut based spatiotemporal regularization provides more consistent activation patterns by regularizing the propagation front and removing false activations outside of the primary visual cortex.



(a) Medial view (b) Zoom of the calcarine fissure

Figure 6.11: Tracking results obtained with visual stimulation of expanding checkerboard rings ($\lambda_{\text{space}} = 3$, $\lambda_{\text{time}} = 1$). Color codes for the first time of activation during the time window considered, between 2310 and 2367 ms after the beginning of the stimulation. Colormap was reduced to 6 colors to present a clearer representation of the propagation. Source estimates were obtained with a spherical forward model and a minimum-norm inversion using unconstrained orientations. Source amplitudes at each position were normalized using prestimulation recordings (see text). Triangulation has about 50 000 vertices. The graph cut based spatiotemporal regularization provides consistent activation patterns by regularizing the propagation front and exhibiting active regions in the primary visual cortex.

6.3.2 Results on somatosensory data

Data acquisition and analysis

Acquisition of the somatosensory data was done with the same MEG device that was used for the visual stimulation paradigm. The somatosensory stimulation was an electrical square-wave pulse delivered randomly to the thumb, index, middle, and little finger of each hand of a healthy right-handed subject. The stimulus intensity was below the motor threshold. In order to improve the SNR, 400 recordings were averaged for each finger. These data come originally from the study presented by S. Meunier *et al*. in [150]. To produce precise tracking results, the triangulation over which cortical activations have been estimated was sampled with a very high number of vertices (about 55 000). The forward modeling was performed with a spherical head model. The source activations were computed with an ℓ_2 prior using constrained orientations. The reason for using constrained orientations with this dataset, is that the shape of the cortical mantle around the brain regions activated by such a stimulation is much simpler than in the occipital region around the calcarine fissure. Dipole orientations provided by the normals to the mesh obtained by segmentation of the gray matter are in this case better estimated.

Somatosensory stimulation is commonly used to validate M/EEG methods since it is known that somesthetic inputs project in precise brain areas [104]. Among these areas are the primary motor cortex (S1) and the secondary motor cortex (S2). While the amplitude of the first activation in S1 is high, the activation that appears later after stimulation in S2 is much weaker but still present. This leads to the conclusion that a same threshold for both activations in S1 and S2 is bound to fail. This issue can be compared to the illumination problem when tracking objects in video sequences. The object keeps moving but its intensity and contrast can change over time. To tackle this problem, the tracking algorithm on the somatosensory dataset requires as preprocessing to construct a particular data cost function

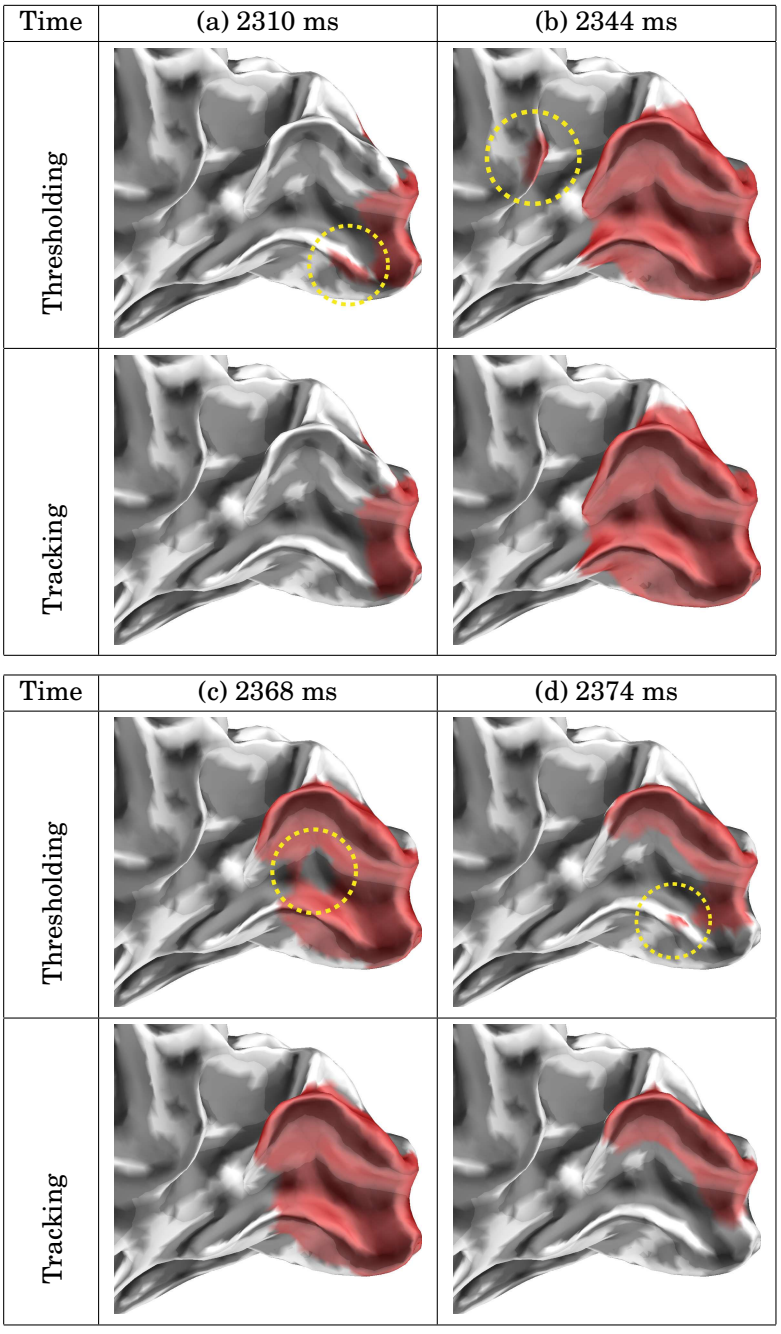


Figure 6.12: Comparison between naive thresholding and tracking with spatiotemporal regularization. Tracking and thresholding results are presented at multiple time instants during visual presentation of the expanding checkerboard rings. Thresholding corresponds to $\lambda_{\text{space}} = 0$ and $\lambda_{\text{time}} = 0$, *i.e.,* no regularization, while the tracking is performed with $\lambda_{\text{space}} = 3$ and $\lambda_{\text{time}} = 1$. (a) illustrates how the tracking manages to remove the spatially inconsistent activation on the lower part of V1. (b) illustrates how the tracking makes the incorrect activation on the anterior part of the parieto-occipital fissure (cf. figure 6.9) disappear. (c) shows how the tracking fills the hole in the active region. This is consistent with the retinotopic organization of V1 and the rings used in the visual stimulation. (d) is another illustration of the regularization, here during the half period when the activation leaves V1.

that enables to define a common threshold at any time after stimulation while still using the essential information that is the source amplitude.

Designing f with heterogeneous activation levels

M/EEG data have typically a sampling rate around 1000 Hz and the characteristic times of the phenomena that are being recorded are about a few milliseconds. Although this time of course depends on the type of neural activations, it suggests that an activity is significant if it lasts for a few consecutive time instants. Hence, the construction of a function f based on small time windows rather than each time instant is still relevant. From now on, k indexes time windows rather than time instants. These windows are denoted by w_k.

The natural idea behind the choice of f is that a vertex is very likely to be active during a time window w_k if its activity is close, in relative distance, to the activation of a source that captures a significant amount of energy.

Let $\mathbf{a}_i^0(k)$ denote the activation time series of vertex i during window w_k (for the time being the superscript 0 can just be ignored). The template $\mathbf{a}_{i_0}^0(k)$ is defined as the time series that captures the highest amount of energy ($i_0 = \arg\max_i \|\mathbf{a}_i^0(k)\|_2$). The function f can now be defined over window w_k by:

$$f_k(i) = \frac{\|\mathbf{a}_i^0(k) - \mathbf{a}_{i_0}^0(k)\|_2}{2\|\mathbf{a}_{i_0}^0(k)\|_2} \in [0,1] \tag{6.7}$$

The function f is designed to take its values between 0 and 1 in order to facilitate the choice of the threshold T. The influence of the activation level does not appear any more directly in f but only in the choice of the template.

It is however possible to use multiple templates since activations are very likely to be simultaneously localized in different regions having different temporal types of activations. This is done using a greedy algorithm similar to standard matching pursuit algorithms. Note that such greedy approaches have been successfully used in the field of M/EEG with the RAP MUSIC inverse problem solver [153]. In both procedures, the most significant source is first estimated. Its contribution to the data is then removed before looking for the next significantly active dipole. This continues until the data have been sufficiently well explained. Although the RAP MUSIC algorithm uses signal subspaces, the idea developed here using temporal templates is fairly similar.

Let $\mathbf{A}_k^0$ be the matrix of all activations during window k, $\mathbf{A}_k^0 = (\mathbf{a}_i^0(k))_i$. Each column of $\mathbf{A}_k^0$ contains the activation time series of one dipole. The objective is to select the best L templates that capture most of the activity during window w_k. The strategy to do so consists in selecting them iteratively. Template l is obtained after template $l-1$ by finding the column of $\mathbf{A}_k^l = (\mathbf{a}_i^l(k))_i$ that has the biggest ℓ_2 norm, $i_l = \arg\max_i \|\mathbf{a}_i^l(k)\|_2$. The matrix $\mathbf{A}_k^{l+1}$ is obtained by projecting the columns of $\mathbf{A}_k^l$ orthogonally to the vector $\mathbf{a}_{i_l}^l$.

$$\mathbf{A}_k^{l+1} = \mathbf{\Pi}_k^l \mathbf{A}_k^l = \left(\mathbf{I} - \frac{\mathbf{a}_{i_l}^l \mathbf{a}_{i_l}^{lT}}{\|\mathbf{a}_{i_l}^l\|_2^2}\right) \mathbf{A}_k^l \tag{6.8}$$

The process ends when $\|\mathbf{A}_k^l\|_2$ becomes smaller than a certain percentage $P \in [0,1]$ of $\|\mathbf{A}_k^0\|_2$ ($\|\ \|_2$ stands for the Frobenius norm). Note that since $\mathbf{\Pi}_k^l$ is a projector, necessarily $\|\mathbf{A}_k^{l+1}\|_2 < \|\mathbf{A}_k^l\|_2$ must hold. When considering multiple templates the function f is defined by:

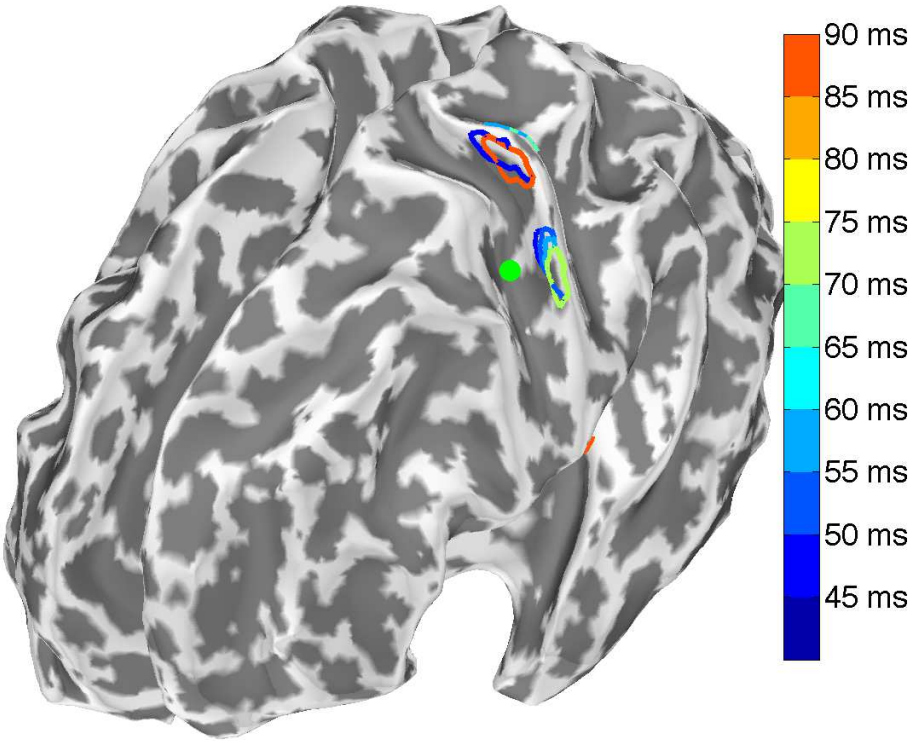
$$f_k(i) = \max_{l=0,\ldots,L-1} \frac{\|\mathbf{a}_i^l(k) - \mathbf{a}_{i_l}^l(k)\|_2}{2\|\mathbf{a}_{i_l}^l(k)\|_2} \in [0,1] \tag{6.9}$$

The procedure has the advantage that the time series $\mathbf{a}_{i_0}^0, \ldots, \mathbf{a}_{i_{L-1}}^0$ correspond to particular

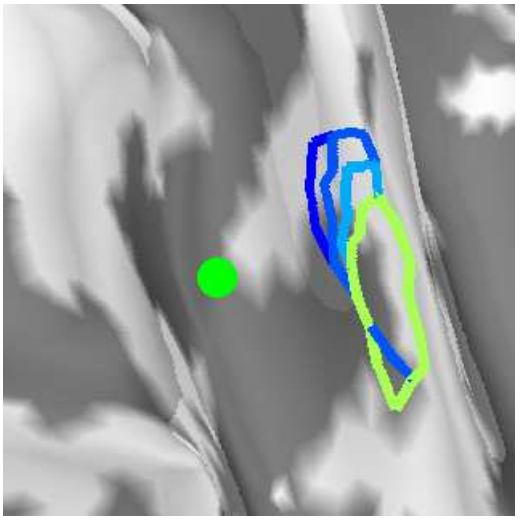
brain locations. More fundamentally, this construction of f has the advantage of making the threshold easy to set. However, during this work several other attempts were also made to define f differently using statistical quantities. Using "noise normalized" inverse methods like dSPM [49] and sLORETA [171], it appeared to be impossible to define a common threshold leading to well localized activations both for the activation peak around 45 ms as well as for later responses. Permutation tests [169] were also investigated using time dependent thresholds computed based on a given p-value. Thresholds obtained using False Discovery Rates (FDR) [85] were also experimented. Both of these approaches, particularly computationally time demanding, did not produce very satisfactory results. The proposed approach has the advantage of speed and simplicity when it comes to selecting the parameters, making the tool quite easy to use.

Tracking

The window size was set equal to 20 ms with an overlap between neighboring windows of 75%, *i.e.,* 15 ms. For each window, f was computed using multiple templates with $P = 0.2$. In practice, a maximum of 3 templates were chosen within each time window. The threshold was set equal to $T = 0.2$. Results with the right index finger are presented in figure 6.13. Color codes for time, or equivalently a window index, and indicates the border of the active region in each time window. This result provides a representation of the brain dynamics after stimulation of the right index finger. Cortical activations are successfully tracked over time, as early as 30 ms after stimulus onset in left primary somatosensory cortex (S1), during the displacement of neural activity along the postcentral gyrus all the way to the secondary somatosensory cortex (left and right) and the left Brodmann area 5. This confirms what is reported in [104] about the processing of such somatosensory tasks by the human brain. In order to illustrate the influence of the spatiotemporal regularization on the solution, results with no regularization at all, and no temporal regularization are provided in figure 6.14. It can be observed that the spatiotemporal regularization is actually required to prune the spurious activations.

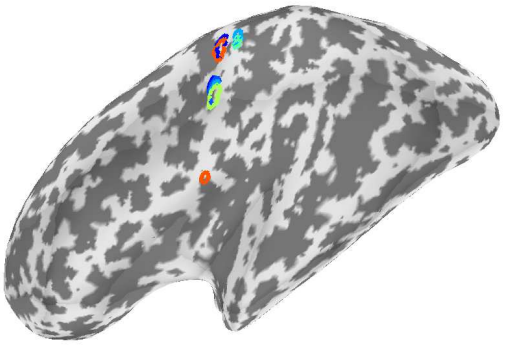


(a) Result on the partially inflated cortex. The green dot corresponds to the location of the equivalent current dipole located at 44 ms after stimulation. Colored lines correspond to the border of the active region at different time instants (1 to 3).

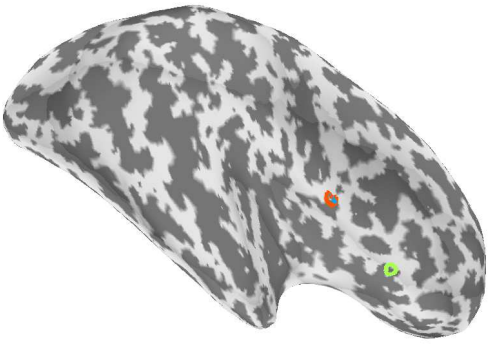


(b) Zoom on the tracking result in S1.

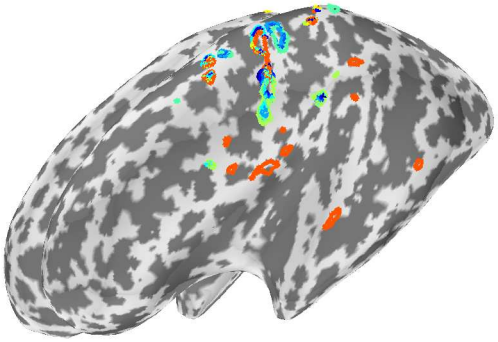
Figure 6.13: Result of tracking using the graph cut algorithm on somatosensory dataset. Source estimates were obtained with a spherical forward model and a minimum-norm inversion on MEG data obtained from a somatosensory evoked response study. Triangulation has about 55 000 vertices. Data presented here are for the stimulation of the right index finger. Cortical activations are tracked over time ($\lambda_{\text{space}} = 0.05$ and $\lambda_{\text{time}} = 0.05$), as early as 30 ms after stimulus onset in left primary somatosensory cortex (S1), during the displacement of neural activity along the postcentral gyrus all the way to the secondary somatosensory cortex (left and right) and left Brodmann area 5 [104].



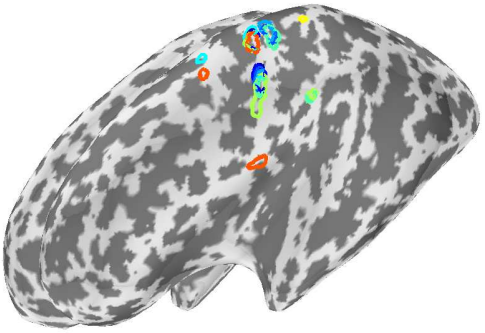
(a) Tracking result on inflated cortex: Left hemisphere



(b) Tracking result on inflated cortex: Right hemisphere



(c) Result with no regularization ($\lambda_{\text{space}} = 0$, $\lambda_{\text{time}} = 0$)



(d) Result with no temporal regularization ($\lambda_{\text{space}} > 0$, $\lambda_{\text{time}} = 0$)

Figure 6.14: Influence on the regularization on the tracking result on the somatosensory dataset. Source estimates were obtained with a spherical forward model and a minimum-norm inversion on MEG data obtained from a somatosensory evoked response study. Triangulation has about 55 000 vertices. Data presented here are for the right index finger. It can be observed that the spatiotemporal regularization is actually required to prune the spurious activations.

6.4 CONLUSION

In this chapter a method has been proposed to address the challenging problem of robustly estimating the spatiotemporal evolution of neural activities rather than just localizing specific brain areas involved in experimental tasks. The approach is based on an optimization over the active domain that penalizes spurious activations and extracts activity with spatiotemporal regularity. A Lagrangian formulation is derived and it is shown how the functional obtained can be discretized over a triangulation and efficiently optimized with a graph cut based procedure.

The principled method proposed here allows the active domain to have topology changes. It implies that active regions can appear, split, merge and disappear during the time window of interest. A source is not necessarily flagged as active during the full time period of interest. Thanks to the use of a very efficient graph cut implementation, with a linear complexity observed in practice, the optimization can be run in a few seconds on real M/EEG datasets with highly discretized cortical meshes.

A possible perspective for this work is to integrate into the temporal regularisation term a prior coming from diffusion MRI data. By measuring white matter anisotropy such data provide insights on subcortical neural pathways. With the latter graph-based formulation such an information could be easily integrated by adding temporal edges between triangles strongly linked by the subcortical fiber bundles present in the white matter. Possibly, such links could be defined between non neighboring time instants in order to model delays in the propagation of the neural activations. Moreover, using the residual capacities of the edges after computing the *min-cut*, it may be possible to quantify which subcortical edges have influenced the solution, leading to fruitful insight about how the processing of the information has been distributed across the different functional brain regions involved in the task.

The source code of the tracking algorithm and the demo scripts necessary to reproduce the figures of this chapter are available in a Matlab toolbox called EMBAL (Electro-Magnetic Brain Activity Localization):

`https://gforge.inria.fr/projects/embal`

CHAPTER 7

GRAPH-BASED ESTIMATION OF 1-D VARIABILITY IN EVENT-RELATED NEURAL RESPONSES

Up to here, much interest has been put on the forward and inverse problems assuming the signal of interest had been accurately captured by the M/EEG sensors. As it is illustrated in this chapter, various factors can cause the estimation of the evoked neural response to be biased.

Classical source estimates are computed from data obtained by averaging many repetitions of recordings measured under the same conditions. The measurements are realigned in time according to a common reference, typically the stimulus onset, and then averaged to get a clean evoked potential (EP). However, by doing so, it is assumed that the evoked neural response stays the same across the different repetitions, a.k.a., trials. This is unfortunately not true. The event related potential (ERP) can for example vary in latency, amplitude or frequency typically because of habituation effects, anticipation strategies, or fatigue of the subject.

In this chapter, we address the challenging problem of single-trial data processing. We propose to make use of recent progress in graph-based methods in order to achieve parameter estimation on single-trial data and therefore limit the estimation bias on the evoked response. The method exposed guarantees global optimality of the solution, hence avoiding initialization problems. Contrary to many alternative methods, it also avoids the use of the average data in the computation and the necessity to define an *a priori* model for the response. The algorithm is data-driven and works in two steps. First, the graph Laplacian offers a convenient way of reordering the dataset with respect to the response latency. And second, the actual estimation of the response latency across trials is performed in a robust way with a graph cuts algorithm. The full processing does not require any manual tuning since a method to automatically set the parameters is also detailed.

Results of a simulation study are presented, demonstrating the ability of the method to handle datasets with low SNR as it the case with M/EEG single-trial data. Results on an EEG auditory oddball dataset are also presented.

Contents

7.1 Introduction

Stimulus-locked averaging is applicable for very early (primary) responses whose characteristics are very stable, such as early somatosensory evoked potentials. However, for later responses, corresponding to a more complex treatment of information by the brain, characteristics such as latencies and amplitudes are usually variable across trials, typically because of habituation effects, anticipation strategies, or fatigue [130]. For responses with latency variability, signals time-locked to the stimulus onset are not properly aligned with respect to the timing of the single-trial brain responses, and this leads to considerable blurring when the averaged evoked response is computed. The event related brain response is hence inaccurately estimated, and this can lead to wrong conclusions about the latencies between neural activations [123]. This chapter provides methodological tools to estimate the variability, *e.g.,* the latency, of brain responses across trials, and thus to correct for the averaging bias.

Analysis of single-trial EEG was pioneered by Woody's cross-correlation averaging [230]. In this work, the latency of single-trial event-related activity is estimated by using a template, supposed to model accurately the evoked response. The algorithm alternately estimates the template and corrects for the latency bias in each trial by finding the maximum of a cross-correlation. In each iteration, the template is updated with the average of the corrected data. Convergence is observed in practice, but not guaranteed. The obtained solution may not be globally optimal since it depends on the time series used to initialize the algorithm. Subsequent work has extended Woody's idea by including amplitude variability and by placing the estimation of single-trial parameters in a maximum likelihood framework [114, 134, 147, 206, 207]. Direct denoising of EEG single-trial data with a time-scale decomposition has been proposed [181], which designs a wavelet template from the average signal across trials. As in [230], the average signal is explicitly used as a template in the computation. Several other methods, based on linear decompositions or wavelet analysis [15, 16, 185, 221], make the assumption that the evoked response can be well represented by a linear combination of functions within a dictionary. The dictionary is provided as a prior to the algorithm or learned from the data, *e.g.,* with a singular value decomposition (SVD).

All methods described above suffer from various pitfalls: the lack of proof of convergence of the procedure, the dependence of the results on the initialization, the use of the average data in the computation, and the necessity to define an *a priori* model for the waveform of the response.

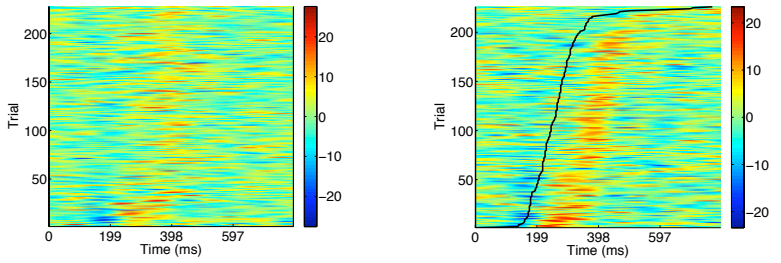
This chapter approaches cross-trial variability from a different perspective, in a data-driven way, free of the pitfalls just enumerated. The problem is cast into an optimization framework in which global optimality can be proved, and where initialization is not an issue. Moreover, the solution can be found very efficiently by using associated fast algorithms. The method proceeds in two stages. First, we propose to sort the trials automatically according to the variability of the evoked potentials without estimating it explicitly. The estimation of the variability is performed in a second stage.

As an illustration, figure 7.1 presents raster plots of a raw dataset (a) and a reordered dataset (b): in these images, each line represents a trial. In figure 7.1(b), the trials have been reordered according to the measured motor response of the subject. A structure thus becomes visible and can be used to interpret the single-trial ERPs. Such representations of multi-trial ERP recordings, called ERP images were pioneered in [123] and made generally available through EEGLAB [59]. EEGLAB proposes several ways of reordering ERP images, according to event triggers, or to the phase of time-frequency decompositions.

In this chapter, we propose a reordering, based only on the evoked response, without relying on any external information. In some multi-trial ERP recordings, it is reasonable to assume that a similar neural activation occurs in each repetition of the experiment, but with

a latency between stimulation and response that is variable across trials. This leads to the intuition that the latency of the response is the main "degree of freedom" within the data. The single-trial time series then lie on a noisy one-dimensional manifold which can be parameterized by this latency. In order to capture this "degree of freedom", manifold-learning techniques are very well-suited: by providing low dimensional representation of the data, they offer an efficient way of revealing the structure present in a dataset. We propose to use methods based on eigendecompositions of graph Laplacians [14, 37]. Nonlinear dimensionality reduction methods have been applied before to functional MRI data [7]. In [8] and [198], low dimensional representations of fMRI datasets have been used to identify and classify brain functional regions. For EEG data, our previous work [94] has shown how Laplacian Eigenmaps [14] can be used to reveal the structure of an EEG dataset.

In this chapter, the nodes of the constructed graphs correspond to single-trial EEG time series and it is shown on an EEG sample dataset that the "best" one dimensional representation of the data obtained with graph Laplacian procedures is monotonically related to the response latency. Equivalently, it is shown that the first coordinate in the low dimensional space can be used to reorder the dataset as in figure 7.1(b). We will refer to this step of the method as the *spectral reordering* of the raster plot.



(a) Original raster plot

(b) Raster plot reordered with the measured motor response of the subject.

Figure 7.1: Illustation of raster plot reordering on real EEG recordings. Each line of the image represents a time series and the color codes for the signal amplitude in μV. In this dataset the measured latencies of the motor responses, represented by the dark semi-vertical line in figure 7.1(b), are used to sort single-trial time series.

Spectral reordering reveals the structure present in the raster plot, but it does not explicitly estimate the trial-dependent parameters in each trial. This problem is tackled in the second step of the procedure, eventually leading to optimized event-related averaging. Given a reordered raster plot, estimating the latencies corresponds to finding an increasing function similar to the one defined with the dark semi-vertical line in figure 7.1(b). This function takes its values on the grid defined by the image, and can only assume a finite number of discrete values. The problem of extracting the latency information from the reordered raster plots can thus be viewed as a combinatorial problem, which can be very efficiently optimized using a graph cut algorithm. This step is performed independently from the spectral reordering step, and only requires a sorted raster plot as input.

The chapter is organized as follows. In section 7.2 we introduce Manifold Learning and present the graph Laplacian method. In section 7.3, we apply it to the reordering of multi trial EEG data and show how the proposed approach compares to the more classical Principal Components Analysis (PCA). Section 7.4 focusses on latency estimation and formulates it as a combinatorial optimization problem. An efficient solution to this problem is proposed by computing a minimal cut on a specially designed graph. Finally, a strategy to estimate the different parameters is detailed and the robustness of the procedure is investigated by

numerical simulations. Results on synthetic and real data accompany each of the methods presented.

7.2 MANIFOLD LEARNING

Let $(x_i)_{i=1,\dots,N}$ be N elements of a metric space $(\mathcal{X}, d_{\mathcal{X}})$, which are distributed with a probability distribution p on a low-dimensional smooth sub-manifold $\mathcal{M}$ of $\mathcal{X}$. In this chapter, the $(x_i)_{i=1,\dots,N}$ are time series and N the number of trials.

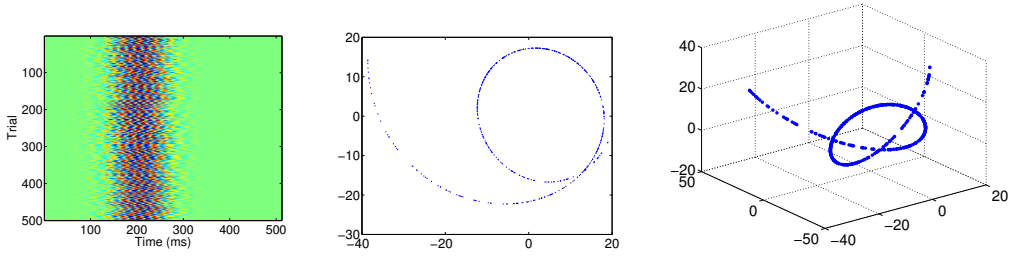
This section deals with manifold learning techniques: Section 7.2.1 presents a Principal Components approach to variability analysis. Sections 7.2.2 and 7.2.3 recall notions on non-linear embedding, leading to the Laplacian embedding algorithm. Differences between the linear and non-linear approaches are emphasized on synthetic datasets in section 7.2.1 and will also be illustrated in section 7.3.1.

7.2.1 Principal Component Analysis

Principal Component Analysis represents the data in a new coordinate system, obtained through a rotation that diagonalizes the empirical covariance matrix. Although PCA per se does not modify the dimensionality of the representation, by ordering the eigenvalues of the empirical covariance matrix, one can represent the data in the leading PCA directions. The PCA representation is a valuable tool for exploratory analysis as it can provide a representation of the structure present in the data.

To take the example of latency variability, consider a dataset of translated versions of a reference template $x_i(t) = x(t - \tau_i)$. The data is wide-sense stationary, and its covariance matrix C_x is diagonalized in the discrete Fourier basis. The Fourier transform of the translated signal $x_i(t) = x(t - \tau_i)$ is phase modulated: $\hat{x}_i(\omega) = e^{i\tau_i\,\omega}\,\hat{x}(\omega)$, where ω denotes the frequency.

Principal Components correspond to the frequencies that dominate the power spectrum of the signal $P_x(\omega)$ (the Fourier transform of the covariance matrix C_x). Consider the dominant frequency ω_1: the coordinates of each trial x_i in the two first PCA directions are $\cos(\tau_i\omega_1 + \phi)|\hat{x}(\omega_1)|$ and $\sin(\tau_i\omega_1 + \phi)|\hat{x}(\omega_1)|$, where ϕ is the phase of $\hat{x}(\omega_1)$. The data thus organize along a one-dimensional manifold parameterized by latency τ_i.



(a) Original raster plot (a set of jittered Gabor wavelets)

(b) 2D PCA projection of the set of translated time-courses

(c) 3D PCA projection of the set of translated time-courses

Figure 7.2: PCA analysis of a set of 500 jittered time series of 512 time samples.

To illustrate this, a set of 500 jittered time series, each with 512 time samples is displayed as a raster plot in Figure 7.2(a). The projections of the data in the leading two and three PCA dimensions, in Figure 7.2, cluster along curves, indicating the 1D structure of the data. Reordering the time-courses with respect to latency is equivalent to finding a parameteriza-

tion of the curves in Figure 7.2(b) and 7.2(c). In the next subsections, we present methods to automatically reorder time series according to their 1D variability.

7.2.2 Nonlinear embedding

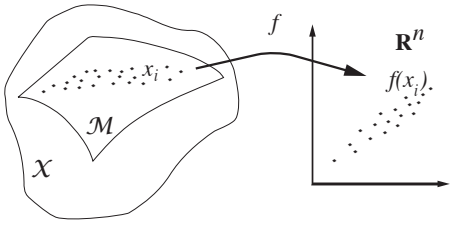


Figure 7.3: Non-linear embedding into a low-dimensional Euclidian space

Given the distance $d_{\mathcal{X}}$ and the points (x_i), the aim is to recover the structure of $\mathcal{M}$ via an embedding function f that maps the (x_i) into a low-dimensional Euclidian space $\mathbb{R}^n$ (cf. figure 7.3). The embedding f provides a low-dimensional representation of the dataset and also a parameterization of the manifold. When $\mathcal{M}$ has a 1-d structure, the first coordinate of f can be used to order the points (*i.e.,* in our context, the time series), provided that the function f satisfies a *"regularity"* constraint: if two points x and z are close in $\mathcal{M}$, then so must $f(x)$ and $f(z)$ in $\mathbb{R}^n$. This is sometimes referred to as a minimal distortion property.

For $n = 1$, a Taylor expansion provides the following inequality [14]:

$$|f(z) - f(x)| \leq d_{\mathcal{M}}(x, z)\|\nabla f(x)\| + o(d_{\mathcal{M}}(x, z)) \tag{7.1}$$

where ∇f stands for the gradient of f and $d_{\mathcal{M}}$ is the geodesic distance on the manifold between points x and z. The notation $g(z) = o(d_{\mathcal{M}}(x, z))$ means that $\frac{g(z)}{d_{\mathcal{M}}(x,z)}$ tends to 0 as z tends towards x. The inequality in equation (7.1) means that $d_{\mathcal{M}}(x, z)\|\nabla f(x)\|$ is a good first order upper bound for $|f(z) - f(x)|$.

In order to obtain an embedding that satisfies the *"regularity"* constraint, Laplacian-based methods try to control the smoothness of f globally by minimising

$$\int_{\mathcal{M}} ||\nabla f(x)||^2 p(x)^s dx \ ,$$

provided that

$$\int_{\mathcal{M}} ||f(x)||^2 p(x)^s dx = 1 \ \ (i).$$

The latter condition removes the scaling indetermination for the function f. The parameter s controls the influence of the probability density on the solution. If s is strictly positive, the regularity of f will be more constrained in high density regions.

The optimization problem under constraint *(i)* can be formulated as a saddle-point problem for the Lagrangian $\mathbf{L}(f, \lambda)$:

$$\mathbf{L}(f, \lambda) = \int_{\mathcal{M}} ||\nabla f(x)||^2 p(x)^s dx + \lambda \left(1 - \int_{\mathcal{M}} ||f(x)||^2 p(x)^s dx\right) \ . \tag{7.2}$$

Introducing the s-weighted Laplacian operator defined as $\Delta_s f \triangleq -\frac{1}{p^s}\mathrm{div}(p^s \nabla f)$ and $\langle f, g\rangle_{\mathcal{M}} \triangleq$

$\int_{\mathcal{M}} \langle f(x), g(x) \rangle p(x)^s dx$, $\mathbf{L}(f, \lambda)$ defined in equation (7.2) can be rewritten:

$$\begin{aligned}
\mathbf{L}(f, \lambda) &= \int_{\mathcal{M}} ||\nabla f(x)||^2 p(x)^s dx + \lambda \left(1 - \int_{\mathcal{M}} ||f(x)||^2 p(x)^s dx\right) \\
&= \int_{\mathcal{M}} \langle \nabla f(x), \nabla f(x) \rangle p(x)^s dx + \lambda \left(1 - \int_{\mathcal{M}} \langle f(x), f(x) \rangle p(x)^s dx\right) \\
&= \int_{\mathcal{M}} \langle p(x)^s \nabla f(x), \nabla f(x) \rangle dx + \lambda \left(1 - \int_{\mathcal{M}} \langle f(x), f(x) \rangle p(x)^s dx\right) \\
&= \int_{\mathcal{M}} -\langle \mathbf{div}(p(x)^s \nabla f(x)), f(x) \rangle dx + \lambda \left(1 - \int_{\mathcal{M}} \langle f(x), f(x) \rangle p(x)^s dx\right) \quad (*) \\
&= \int_{\mathcal{M}} \langle \Delta_s f(x), f(x) \rangle p(x)^s dx + \lambda \left(1 - \int_{\mathcal{M}} \langle f(x), f(x) \rangle p(x)^s dx\right) \\
&= \langle \Delta_s f, f \rangle_{\mathcal{M}} + \lambda \left(1 - \langle f, f \rangle_{\mathcal{M}}\right)
\end{aligned}$$

Step (*) comes from the fact that the gradient operator is the negative adjoint of the divergence.

Differentiating $\mathbf{L}(f, \lambda)$ with respect to f leads to:

$$\frac{\partial \mathbf{L}(f, \lambda)}{\partial f} = \Delta_s f - \lambda f$$

Therefore setting this derivative to zero imposes f to satisfy:

$$\Delta_s f = \lambda f$$

thus f to be an eigenvector of the operator Δ_s associated to the eigenvalue λ.

Notice that if $\Delta_s f_i = \lambda_i f_i$,

$$\mathbf{L}(f_i, \lambda_i) = \lambda_i$$

This implies that optimizing the constrained problem whose Lagrangian is given by $\mathbf{L}(f, \lambda)$ requires to find the eigenvectors of Δ_s with the smallest eigenvalues.

The constant function f_{cst} equal to 1 everywhere is an eigenfunction of Δ_s for the eigenvalue 0. To avoid this trivial embedding, the solution f is constrained to be orthogonal to the function f_{cst}, *i.e.,* $\int_{\mathcal{M}} f(x) p(x)^s dx = 0$ (ii).

The optimal embedding f under constraints *(i)* and *(ii)* is the eigenfunction of Δ_s corresponding to the smallest non-zero eigenvalue.

In order to compute f from a manifold sampled with a limited number of points, the operator Δ_s needs to be approximated. Graph Laplacian methods approximate Δ_s by the Laplacian of a particularly designed graph. Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be an undirected graph where $\mathcal{V}$ are the nodes $(x_i)_{i=1,\dots,N}$ and $\mathcal{E}$ are the edges. A weight w_{ij} is associated to every edge $(i, j) \in \mathcal{E}$, leading to a weighted graph $\mathcal{G}$. The Laplacian L of the graph is a matrix defined by $L = D - W$ where $W = (w_{ij})_{ij}$ and D is diagonal with $D_{ii} = \sum_j w_{ij}$.

The random-walk Laplacian is a normalized version of L defined by $L_{rw} = D^{-1}L = I - D^{-1}W$ where I is the identity.

A weighting matrix W yielding a good approximation of Δ_s must now be defined. This is done with the help of a similarity measure k which is a non-increasing function of the distance $d_{\mathcal{X}}$. Here, k is a Gaussian kernel with standard deviation σ:

$$k(x_i, x_j) = e^{-\frac{d_{\mathcal{X}}(x_i, x_j)^2}{\sigma^2}} \ .$$

Results of [107] show that, for a given k, the random-walk Laplacian of $\mathcal{G}$ converges almost

surely to Δ_s when the sample size N goes to infinity, if

$$w_{ij} = \frac{k(x_i, x_j)}{(d(x_i)d(x_j))^{1-s/2}} \tag{7.3}$$

where $d(x) = \sum_{i=1}^{N} k(x, x_i)$ is an estimator of the probability density function over $\mathcal{M}$, The embedding function can therefore be obtained by computing the eigenvectors f_q of L_{rw} satisfying:

$$\begin{aligned} (I - D^{-1}W)f_q &= \lambda_q f_q \\ (D - W)f_q &= \lambda_q D f_q \\ Lf_q &= \lambda_q D f_q \end{aligned} \tag{7.4}$$

It can easily be proved that 0 is a trivial eigenvalue and also that L is a symmetric positive matrix, which implies, since $d(x_i) \geq 0$ for all i, that the generalized eigenvalues λ_q are all positive and can be ordered:

$$0 = \lambda_0 \leq \lambda_1 \leq \cdots \leq \lambda_q \leq \lambda_{q+1} \leq \cdots \leq \lambda_{N-1}.$$

The embedding f into $\mathbb{R}^n$ is then given by:

$$f(x_i) = (f_1(i), f_2(i), ..., f_n(i)).$$

where $f_q(i)$ is the ith component of f_q.

7.2.3 Laplacian embedding algorithm

Observe that, in the case of a uniform sampling over the manifold, the parameter s has no influence. Numerical experiments on synthetic and real EEG data showed that s did not have much influence on the results, and as a consequence s was set to 1. This corresponds to the *Diffusion Map* algorithm [37]. The following algorithm details the different steps to compute.

Laplacian-based n-dimensional embedding:

- Set $d_{\mathcal{X}}$, σ and n the dimension of the embedding.
- Compute K with $K(i,j) = e^{-\frac{d_{\mathcal{X}}(x_i, x_j)}{\sigma^2}}$.
- Compute D_K with $D_K(i,i) = \sum_{j=1}^{N} K(i,j)$ and $D_K(i,j) = 0$ if $i \neq j$.
- Compute $W = D_K^{-1/2} K D_K^{-1/2}$ (equation (7.3) with $s = 1$)
- Compute D with $D(i,i) = \sum_{j=1}^{N} W(i,j)$ and $D(i,j) = 0$ if $i \neq j$.
- Find the $n+1$ first generalized eigenvectors f_k solution of $(D - W)f_k = \lambda_k D f_k$, $k = 0, \ldots, n$.
- The coordinates of point x_i in $\mathbb{R}^n$ are $(f_1(i), f_2(i), \ldots, f_n(i))$.

Since L is symmetric, its first eigenvectors can be computed efficiently with an iterative method, for example an Implicitly Restarted Arnoldi Method [137]. For computational efficiency, it is possible to set to 0 the w_{ij} below a threshold, leading to sparse matrices, and reducing the computational cost of matrix-vector multiplications at each iteration. Note should be taken that too high a threshold leads to a very coarse approximation of the solution.

In order to provide more insight on the discrete solution, it can be observed from equation (7.4) that the solution f_1 also solves the following optimization problem:

$$\underset{f^T D f=1,\ f^T D 1=0}{\arg\min} f^T L f$$

and that expanding $f^T L f$ gives:

$$f^T L f = \frac{1}{2} \sum_{(i,j)\in\mathcal{N}} w_{ij}(f(i) - f(j))^2$$

Since each term of the sum is positive, minimizing $f^T L f$ requires to minimize each $w_{ij}(f(i) - f(j))^2$ which implies that if w_{ij} is big, *i.e.,* $d_{\mathcal{X}}(x_i, x_j)$ is small, $(f(i) - f(j))^2$ should be small. This can be directly related to the "regularity" constraint mentioned above. Although this comment helps to bridge the gap between the discrete and the continuous formulations of the problem, it can be noticed that, contrary to the continuous formulation, the discrete problem does not provide much insight into the influence of the density p.

Manifold learning methods are "data driven". They capture the structure of the dataset, provided that the chosen distance $d_{\mathcal{X}}$ is appropriate. When dealing with time series, many distance functions can be used. In practice, $d_{\mathcal{X}}$ does not need to be an actual distance, and instead it can measure the difference between features of interest for two elements of the dataset. One may even design $d_{\mathcal{X}}$ to be *blind* to some features of the data, which are irrelevant for the application at hand.

7.3 Spectral reordering of EEG times series

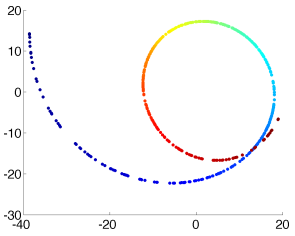
In this context, each x_i is a time series, and $\mathcal{X} = \mathbb{R}^T$ where T is the number of time samples. The same stimulus has been delivered N times leading to N time series.

7.3.1 Toy examples

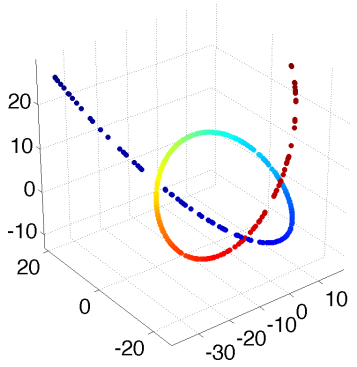
Once computed, the Laplacian embedding f provides a parameterization of the manifold $\mathcal{M}$, which means that if $\mathcal{M}$ has a noisy 1D structure, the first coordinate, f_1, orders the elements along the manifold. This is now illustrated with the noiseless synthetic dataset already presented in Section 7.2.1. It is simulated with $T = 512$ and $N = 500$ points. The embedding was performed using the Euclidian distance and a Gaussian kernel. The embedded points are represented in figure 7.4(c). It can be observed that the embedding unfolds the manifold structure. The ordering provided by f_1 can be encoded with a color, hence each point of the PCA representation can be colored. The 2D and 3D PCA point clouds are presented in figure 7.4(a) and figure 7.4(b): observe that the color changes continuously along the one dimensional structure. Figure 7.4(d) presents the reordering of the raster plot already presented in figure 7.2(a).

To illustrate that the manifold learning method can capture more than a variability in latency, a sample dataset with a variability in latency and scale has been designed (time series are "stretched"). The reordered dataset and the 3D PCA colored embedding are presented in figure 7.5.

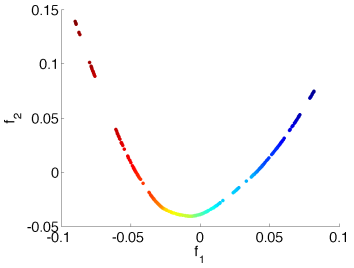
The EEGLAB toolbox offers an alternative way of ordering signals, based on the phase of a time-frequency decomposition using Gabor wavelets. The user is asked to provide the latency of the response, the number of oscillations and the frequency of interest (the frequency



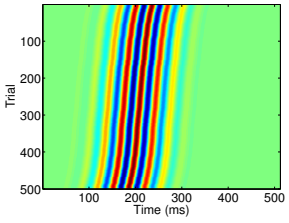
(a) Points of the 2D PCA projection, colored with respect to their first Laplacian embedding coordinate



(b) Points of the 3D PCA projection, colored with respect to their first Laplacian embedding coordinate

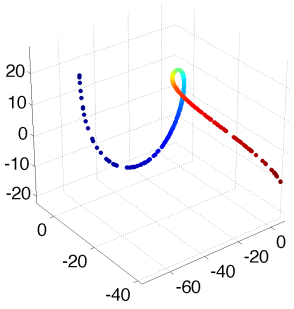


(c) Embedded points colored with respect to the first Laplacian embedding coordinate.

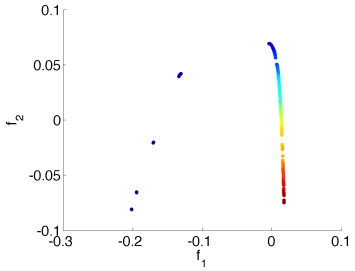


(d) Reordered raster plot

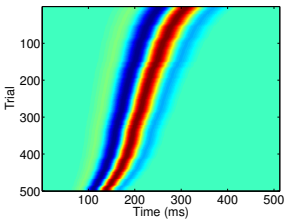
Figure 7.4: Illustration of manifold learning using graph Laplacian on the synthetic dataset of Figure 7.2.



(a) Points of the 3D PCA projection, colored with respect to their first Laplacian embedding coordinate

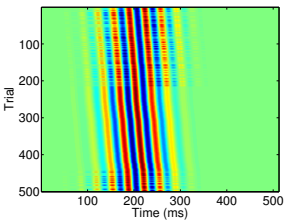


(b) Embedded points colored with respect to the first Laplacian embedding coordinate.

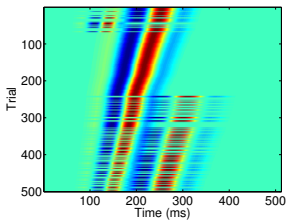


(c) Reordered raster plot

Figure 7.5: Illustration of manifold learning using graph Laplacian on a synthetic dataset with latency and scale variability (time series are "stretched" in time with the increase of the latency).



(a) Phase reordering on the dataset in figure 7.2



(b) Phase reordering on the dataset in figure 7.5

Figure 7.6: Reordering results obtained on the datasets in figure 7.2 and figure 7.5 with the EEGLAB reordering method based on the phase of a Gabor wavelet, with *a priori* defined latency and frequency.

can also be automatically set as maximizing the Fourier spectrum). Once these parameters are set, the phase for the corresponding Gabor function is computed and the raster plot is reordered accordingly. The two datasets presented in figure 7.2 and figure 7.5 are shown, reordered with EEGLAB, in figure 7.6. It can be noticed that the method fails to accurately reorder the raster plots, for different reasons in the two examples. In the first one, the Gabor waveform that has been translated has multiple cycles which implies that the phase in $[0, 2\pi]$ cannot capture the variability. In the second one, due to the temporal scaling, the frequency slightly varies across trials. The EEGLAB procedure assumes a constant frequency across trials, and estimating the phase when the frequency is not well defined leads to errors.

7.3.2 Spectral reordering with realistic time series

The manifold learning method has just been illustrated on synthetic toy examples. We now apply it to a more challenging synthetic problem where the time series are corrupted by an additive noise, and to real ERP recordings.

As a preprocessing, the time series are centered and normalized so that $\|x_i\| = 1$. After this normalization, the half Euclidian distance is given by $\|x_i - x_j\|_2/2 = \sqrt{(1 - \langle x_i, x_j \rangle)/2}$, which implies that it is equivalent to considering the correlation between time series.

As mentioned earlier $d_{\mathcal{X}}$ may not be a real distance but just a similarity measure. To investigate the use of different similarity measures, we propose to use:

$$d_{\mathcal{X};r}(x_i, x_j) = \left(\frac{1 - \langle x_i, x_j \rangle}{2} \right)^{r/2} \tag{7.5}$$

where r is a power exponent applied to the classical Euclidian distance.

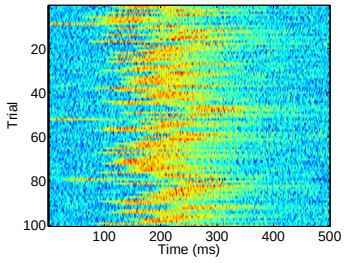
Such a similarity measure is sensitive to time shifting, and is therefore appropriate for capturing the latency information. It also has the advantage that $d_{\mathcal{X};r}(x_i, x_j) \in [0, 1]$ which constrains the choice of σ (see Section 7.5).

The parameters used by the spectral embedding are s (fixed to 1 in practice), r and σ. Section 7.5.1 provides a strategy to set these parameters.

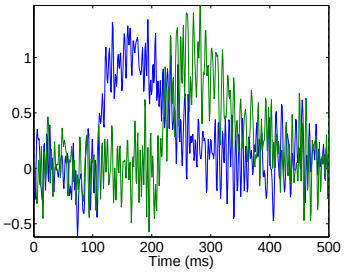
The synthetic data consists in 100 time series ($N = 100$ and $T = 500$) computed from a template (cf. figure 7.11(e), green curve) by translating the positive deflection with a random time lag (with a Gaussian probability distribution of standard deviation σ_{lag}), and adding noise with a SNR measured in variance ratio (variance of signal divided by variance of the noise). Figure 7.7(a), presents a dataset with $\sigma_{lag} = 50$ ms and an additive autoregressive (AR) noise (SNR=1.5=1.76dB). The AR model of order 8 was fitted on spontaneous EEG activity.

The second application of the method is to auditory oddball EEG data. This paradigm

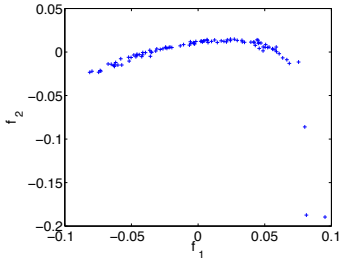
consists of alternating frequent tones and rare ("target") tones. It is known to elicit a positive EEG deflection to the rare tones, referred to as the "P300" or "P3" wave, more prominent on the midline electrodes and occurring at a latency around 300 ms [17]. The data is recorded from the central electrode *Cz* (cf. figure 1.22 in chapter 1), sampled at a rate of 256 Hz and processed with a high-pass filter at 0.5 Hz (Butterworth Zero Phase Filter, Time constant 0.3183 s, 12 dB/oct) and a low-pass filter at 8 Hz (Butterworth Zero Phase Filter, 48 dB/oct). The positive deflection of the P300 wave, in the 3-5 Hz range, is preserved. Figures 7.7(a) and 7.8(a) present raster plots of both data sets. The random nature of the time latency of the P300, as first observed in [130], is obvious in the oddball raster plot.



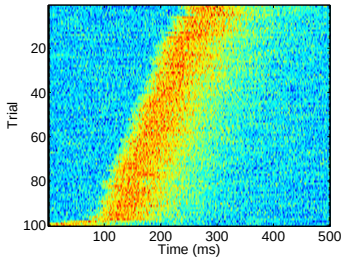
(a) Raster plot of raw time series.



(b) Two time series belonging to the dataset.



(c) 2D embedding of the time series.



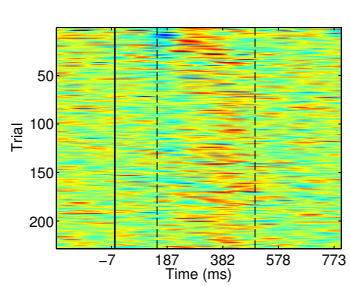
(d) Reordered raster plot of time series using of the first coordinate of the embedding f.

Figure 7.7: Spectral reordering results on synthetic data ($\sigma_{lag} = 50$ ms). Embedding was performed with $r = 2$ and $\sigma = 0.1$.

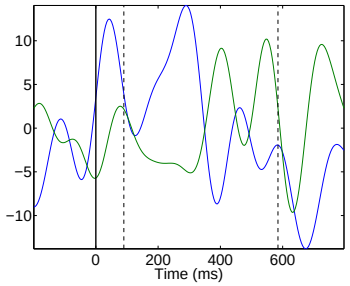
In both cases, the time series were embedded into a two dimensional space (cf. figure 7.7(c) and figure 7.8(c)). In both figures, it can be noticed that the points are clustered along an elongated 1D structure, as it was the case with the toy example in figure 7.4. The first coordinate can therefore be used to correctly parameterize the manifold and to order the time series. By observing the reordered raster plots in figure 7.7(d) and figure 7.8(d), and comparing them to a raster plot reordered using the externally measured motor response, it appears that the first coordinate of f has correctly captured the latency information.

7.4 Robust latency estimation via discrete optimization

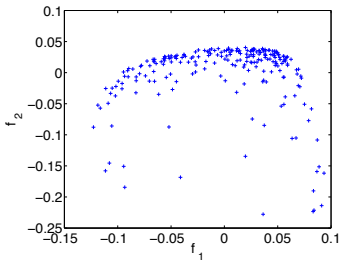
Once the raster plot has been correctly reordered, estimating the latency of the response consists in tracing a non decreasing line through the extrema of the raster plot, like the dark semi-vertical line in figure 7.1(b). Though this may appear a simple task, it is non trivial to



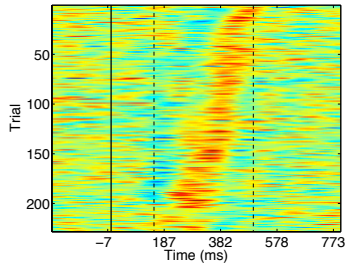
(a) Raster plot of raw time series



(b) Example of time series belonging to the dataset.



(c) 2D embedding of the time series.



(d) Reordered raster plot of time series using the first coordinate of the embedding f.

Figure 7.8: Spectral reordering results on EEG oddball time series. The solid vertical line in the raster plots corresponds to the stimulus onset. The vertical dashed lines provide the limits on the time window used to reorder the data. Embedding was performed with $r = 2$ and $\sigma = 0.05$.

automatize.

In order to automatically extract the latency corresponding to each time series from a reordered plot, one should exploit the global information in the data and take into account the a priori information that latencies are monotonically related in the reordered trials. Finding an increasing function such as the one defined by the dark semi-vertical line in figure 7.1(b), is equivalent to partitioning the raster plot in two, *i.e.,* achieving a two class segmentation of the ERP image, while taking into account the monotony constraint (cf. figure 7.9(b)).

7.4.1 Optimization framework

After reordering, it can be assumed that the latencies of brain responses of two neighboring trials x_i and x_{i+1} are close. Let us denote l_i the latency of the response for trial i. With the Markov Random Field (MRF) optimization framework pioneered in [84], the robust estimation of the $(l_i)_{i=1,\dots,N}$ amounts to solving:

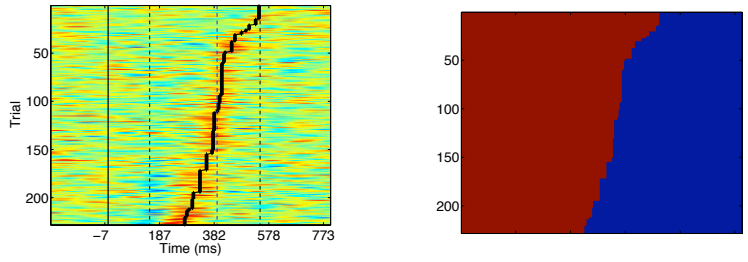
$$(l_i)^*_{i=1,\dots,N} = \underset{(l_i)_{i=1,\dots,N}}{\arg\min}\ E(l_i) \quad \text{with} \quad E(l_i) = \sum_{i=1}^{N} D_i(l_i) + \alpha \sum_{i=1}^{N-1} V_i(l_i, l_{i+1}) \tag{7.6}$$

Each term $D_i(l_i)$, usually called a *data term*, tends to set l_i to the best latency value for time series x_i, regardless of the other series. When considering EEG evoked potentials, high deflections of the signal from the baseline are of particular interest. Thus a reasonable choice is to set l_i so that $x_i(l_i)$ is maximal (or minimal). In our optimization formulation, this leads to $D_i(l_i) = \phi(x_i(l_i))$, where ϕ is a positive decreasing function[1]. In practice, when looking for positive deflections ϕ is defined by $\phi(x_i(j)) = M - x_i(j) \geq 0$ where $M = \max_{i,j} x_i(j)$.

Since single-trial EEG recordings are heavily corrupted by noise, robust estimation of lags is obtained by adding a second term to the energy E. This has been made possible by the spectral reordering, that guarantees that for any i, the time series x_i and x_{i+1} are similar, and thus, neighboring latencies l^*_i and l^*_{i+1} should be close. The *smoothing term* $V_i(l_i, l_{i+1})$ must increase when the distance between l_i and l_{i+1} increases. Without any prior information, it is natural to set all the V_i to a unique function V that does not depend on i.

In our case, we choose $V(l_i, l_{i+1}) = |l_{i+1} - l_i|$ while strongly imposing $l_i \leq l_{i+1}, \forall i$. Parameter α penalizes the difference of latencies between two neighboring time series.

7.4.2 Graph Cuts algorithm



(a) Automatic lag extraction result

(b) Corresponding binary segmentation

Figure 7.9: Result of binary partitioning using the graph cut algorithm applied on the raster plot reordered with motor response (cf. figure 7.1(b))

[1] $D_i(l_i)$ should be positive for technical reasons imposed by the Graph Cuts method – see Section 7.4.2.

The unknown variables l_i take their values on the same time samples as the time series x_i, and can thus only take a finite number of discrete values. This implies that the problem necessarily has a solution, which may however not be unique. In order to find a fast solution of such discrete combinatorial optimization problems, graph-based approaches are being extensively used by the Computer Vision community [23].

MRF (cf. equation (7.6)) optimization with graph cuts is applicable depending both on the discrete set of possible values for l_i and on the choice of the smoothing terms V. They can yield global or local minima. When l_i can take only two values, *i.e.,* labels, global optimality is guaranteed [96]. With multiple labels, global optimality is possible when the l_i can be linearly ordered [23]. This is the case with the one-dimensional problem addressed in this contribution. The optimal solution is here obtained by computing only one *minimal cut* of a specially designed graph. More generally graph cuts methods are adapted for MRF optimization when the smoothing terms lead to a sub-modular energy [127]. See Appendix B for a short introduction to graph cuts.

To construct the graph in our case, it can be observed that a set of latencies satisfying the constraints can be defined by partitioning the raster plot into two classes. This is illustrated in figure 7.9(b) where the optimization has been performed on the raster plot reordered by the motor response figure 7.9(a). The border between the two regions provides an estimation of the latency. Our problem thus amounts to designing a weighted graph that provides a natural correspondence between the partitioning and the energy detailed in equation (7.6).

With the previously defined V, a graph can be constructed as in figure 7.10. Each node of the graph $n_{i,j}$ is indexed by trial number and by time, *i.e.,* a line index i and a column index j, except for the terminal nodes ("source" $\mathcal{S}$ and "sink" $\mathcal{T}$). Cutting the graph in two consists in separating the "source" from the "sink". Edge weights are detailed in table 7.1. Edges with ∞ weights are used to guarantee the constraints: horizontal ∞ weighted edges guarantee that the cut goes through each line once, *i.e.,* l_i is unique for each time series, while vertical ∞ weighted edges guarantee that the solution is increasing.

One can notice that the cost associated to a cut, defined as the sum of the edge weights along the path of the cut, is equal to an energy value. The minimum cut thus provides the optimum. The solution is obtained by a single binary cut, which makes the algorithm extremely fast (a few milliseconds) and also globally optimal. Graph partitioning via minimum cut is, in turn, known to be equivalent to a polynomial problem: the *max flow problem* [73]. We solve the minimum cut problem with the *max flow* algorithm described in [22][2]. The complexity observed in practice is linear in the number of nodes in the graph, *i.e.,* $O(NT)$.

The spectral reordering may provide a raster plot ordered from large latencies to small latencies, in which case one should seek a non increasing partioning of the raster plot. This is why, in practice, the graph cut algorithm is run twice on each dataset, once with the order provided by the manifold learning algorithm and once with this order inverted. The order that leads to the smallest energy is kept.

7.4.3 Result of single-trial latency extraction

The optimization procedure previously described was applied to the reordered datasets displayed in figure 7.7(d) and figure 7.8(d). Results are presented in figure 7.11(a) and figure 7.11(b). Such a lag extraction technique would have been inapplicable on the unordered time series displayed in the raster plots of figure 7.7(a) and figure 7.8(a), and is only made possible after reordering by the non linear embedding.

Figures 7.11(c) and 7.11(d) present the synthetic and the oddball datasets after realignment, *i.e.,* after raster plot reordering and lag correction. The evoked potentials were computed by standard averaging with and without latency correction cf. figure 7.11(e) and fig-

[2]using the open source implementation http://www.adastral.ucl.ac.uk/~vladkolm/software.html

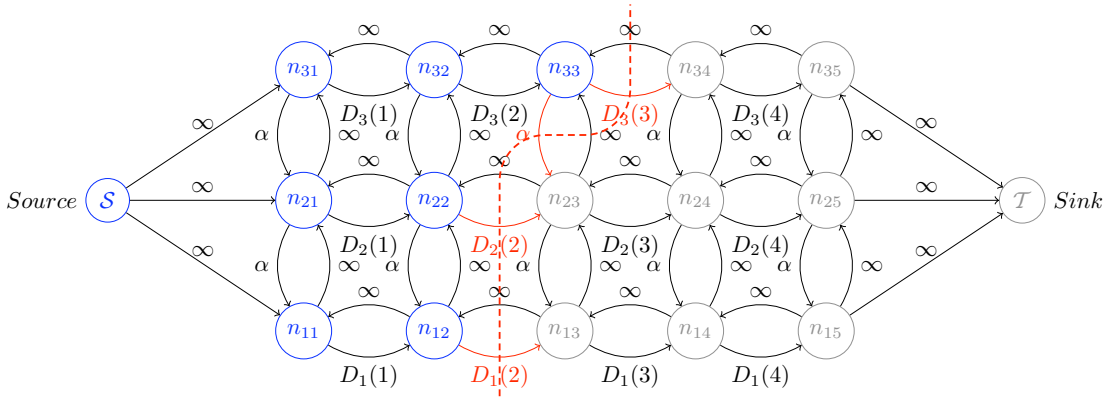


Figure 7.10: Graph illustration for an image $N \times T$ ($N = 3$ time series of length $T = 4$) with an example of minimal cut in red. Time instants index the horizontal edges. In blue are the nodes linked to the "source" node and in gray the nodes linked to the "sink" node after cutting the graph. The corresponding energy is the sum of the weights on the edges (in red) crossed by the minimum cut.

T-Links	Weight
$\mathcal{S} \rightarrow n_{i,1}$	∞
$n_{i,T+1} \rightarrow \mathcal{T}$	∞

N-Links	Weight
$n_{i,j} \rightarrow n_{i+1,j}$	∞ (1)
$n_{i,j} \leftarrow n_{i+1,j}$	α
$n_{i,j} \rightarrow n_{i,j+1}$	$D_i(j) = \phi(x_i(j))$
$n_{i,j} \leftarrow n_{i,j+1}$	∞ (2)

Table 7.1: Edge weights, *i.e.,* link capacities, of the graph for robust time delay estimation. Graph nodes $n_{i,j}$ are indexed by a line index i, trial number, and a column index j corresponding to time. The infinite weight (1) guarantees that l_i^* is unique for i fixed, *i.e.,* that the cut intersects each line only once and (2) guarantees that $l_{i+1}^* \geq l_i^*$, *i.e.,* that the function is increasing. In practice $\phi(x_i(j)) = M - x_i(j) \geq 0$ where $M = \max_{i,j} x_i(j)$.

ure 7.11(f). In the synthetic case, we observe a very good match between the average of the realigned time series and the reference used to generate the data. As expected, realigning the time series provides higher and narrower deflections, because it greatly reduces the blurring effect caused by the variable delays of the neural response.

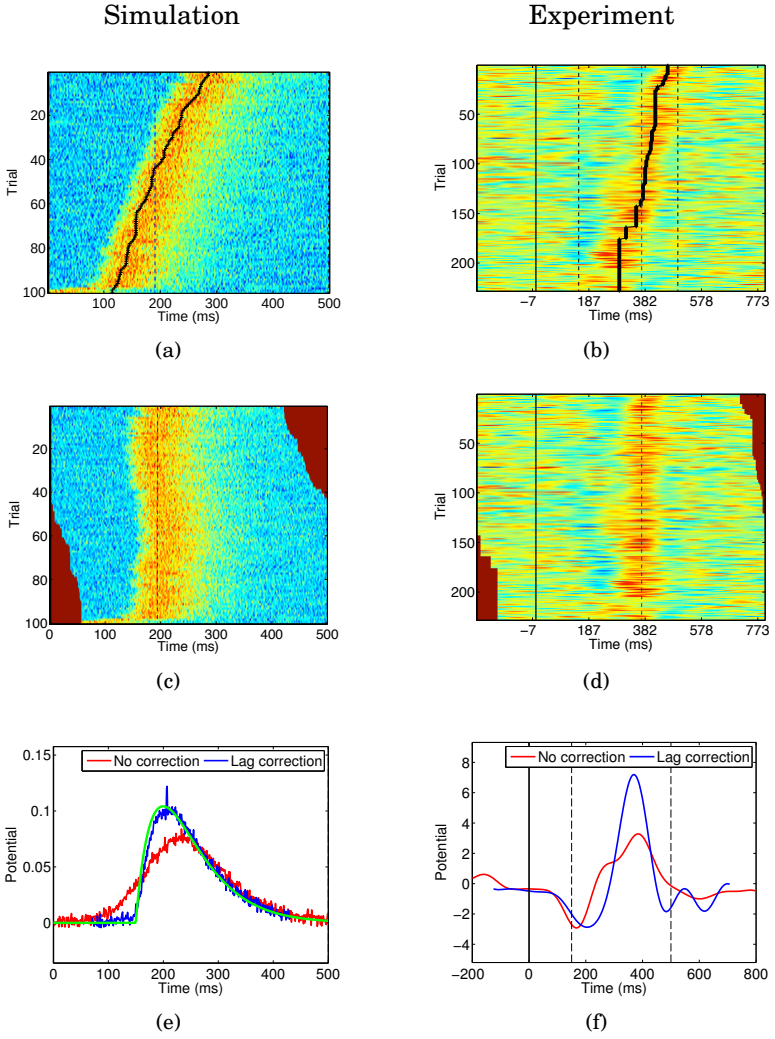


Figure 7.11: Evoked potentials illustrations using single-trial latency estimation. Lags extraction on reordered raster plot of synthetic time series (a) and oddball time series (b). Lags correction on raster plot of synthetic time series (c) and oddball time series (d). (e) evoked potential computed after reordering and lags correction on synthetic dataset (the green curve is the template used to generate the data, the red curve is the evoked potential without reordering and the blue curve is the evoked potential after reordering). We observe a good fit between the template and the evoked potential after reordering. (f) evoked potential computed after reordering and lags correction on the auditory oddball dataset (the red curve is the evoked potential without reordering and the blue curve is the evoked potential after reordering).

7.5 Parameter estimation and robustness

The output of the spectral reordering depends on the definition of the distance $d_{\mathcal{X}}$ and on

the σ used in the Gaussian kernel. It also depends on the parameter s used to weight the Laplacian with the density, although our experience is that the influence of s on the results is negligible. In order to limit the computation time of the following procedure, s was set to 1 by default, corresponding to the *Diffusion Map* algorithm proposed in [37]. According to our definition (7.5), $d_{\mathcal{X}}$ depends on the exponent r. Once the reordering is done, the graph cut method requires the setting of α used to control the regularity of the cut. We propose an automatic way to set these parameters. The robustness of the procedure is evaluated with numerical simulations.

7.5.1 Parameter estimation

The spectral reordering is good if it succeeds in exhibiting the monotonic structure observed for example in figure 7.11(a) and figure 7.11(b). With a proper reordering, the graph cut procedure, constrained to provide non decreasing cuts, reaches lower energy levels. This observation motivates the following strategy to estimate the parameters of the spectral reordering.

Estimating the parameters of the spectral reordering

Let us denote the parameters of the spectral reordering $\theta = (r, \sigma)$ and $E^*_\alpha(\theta)$ the value of the energy reached by the graph cut algorithm with α fixed. The best θ^* is simply obtained by:

$$\theta^* = \arg\min_\theta E^*_\alpha(\theta) \ .$$

In practice, the x_i are normalized and $d_{\mathcal{X};r}$ is designed to take its values in $[0, 1]$. For r and s fixed, the tested values for σ were $(0.1 + 0.1k)_k$ with $k = 1, \ldots, 20$. Using a fixed range of values for σ is made possible by the constraint $d_{\mathcal{X};r} \in [0, 1]$.

For a standard M/EEG dataset with a few hundreds of trials, testing all the different sets of θ takes a few seconds. Lag extraction results for the oddball data are presented in figure 7.13 for different values of α. For α set to 0.1, E_α is displayed in figure 7.12 as a function of σ with $r = 1$ and $r = 2$. The optimum is reached for $r = 2$ and $\alpha = 0.05$.

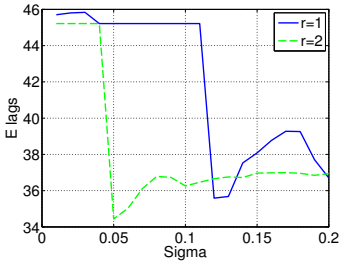
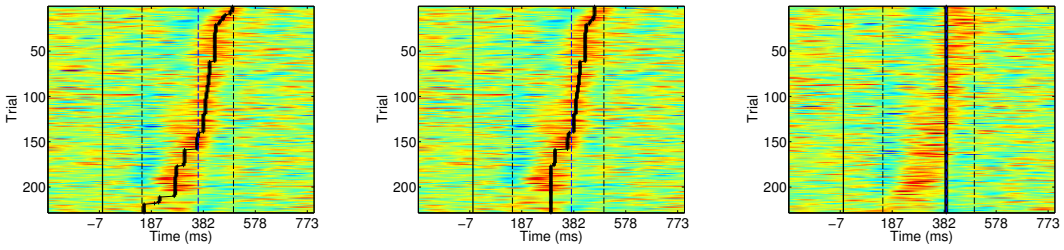


Figure 7.12: E^*_α as a function of r and σ. Computation is done on the oddball dataset with $\alpha = 0.1$.

Estimating α

For α fixed, the above procedure provides a way to estimate the parameters of the spectral reordering. In order to find the α, we propose a method based on K-fold cross-validation.

In figure 7.11(e), it can be observed that the new evoked potential obtained after lag correction matches the template used in the simulation. This validates the procedure in the



(a) $\alpha = 0.01$, $\sigma = 0.05$ (b) $\alpha = 0.1$, $\sigma = 0.05$ (c) $\alpha = 1$, $\sigma = 0.15$

Figure 7.13: Reordered raster plots with lags estimate for different values of α. The σ are automatically defined for α fixed as described in section 7.5.1. Results are obtained with $r = 2$.

synthetic case. Unfortunately, in practice there is no "ground truth". To circumvent this obstacle, the proposed strategy consists in estimating the new evoked potential $e(t)$ on a portion of the data, called the *learning set*, and checking if it correlates well with the rest of the data called the *test set*. The dataset is in practice partitioned into K disjoint subsets $(\mathcal{C}_k)_k$. For $k \in [1, K]$, the test set is the kth subset and the learning set is the rest. With $K = 10$, the evoked potential $e_k(t)$ is estimated on 90% of the data and tested on the remaining 10%.

Since the lag is unknown for the test set, the correlation between the evoked potential and each time series is given by the maximum of the cross correlation with arbitrary lag. The score of generalization S_k is given by:

$$S_k = \sum_{i \in \mathcal{C}_k} \max_{\tau} \langle e_k(\cdot), x_i(\cdot - \tau) \rangle$$

The procedure is run K times leading to a score $S = \sum_k S_k$.

The α that achieves the maximum score S is selected.

In practice the tested values for α were $0, 0.001, 0.01$ and 0.1. As expected, with no noise, the procedure automatically sets α to 0. On the oddball EEG dataset, the procedure leads to $\alpha = 0.1$. This also confirms that using *smoothing terms* $V_i(l_i, l_{i+1})$ is mandatory in the presence noise to obtain a proper estimate of the latencies and subsequently of the evoked response.

With 4 values for α and 10 values for σ, the computation time on the oddball dataset is around 5 minutes per value of r. About half of the time is spent in the computation of the eigenvectors. Note that for each set of parameters the computation is independent. This allows easy parallel computing which could speed up the parameter estimation procedure.

7.5.2 Validation

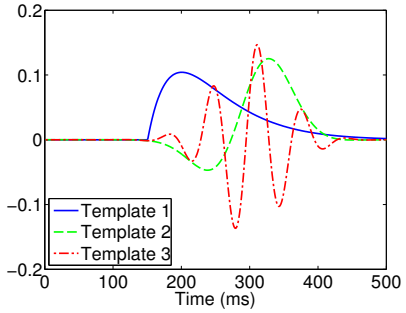
In order to validate the procedure and also investigate its robustness to noise, various numerical experiments were performed on simulated datasets. For each dataset, the parameters were automatically set using the procedure described above. The resulting evoked potential is denoted $e^*(t)$. Assuming the latencies are known, by realigning the data and averaging, one obtains the best evoked potential given the SNR of the dataset. This evoked potential, obtained with the known real latencies, is denoted $e_{rl}(t)$. The real template used to generate the synthetic dataset is denoted $e_{ref}(t)$. The error on the solution was then computed as:

$$\text{Error} = \left| \langle \frac{e^*}{\|e^*\|}, \frac{e_{ref}}{\|e_{ref}\|} \rangle - \langle \frac{e_{rl}}{\|e_{rl}\|}, \frac{e_{ref}}{\|e_{ref}\|} \rangle \right| \in [0, 2] \tag{7.7}$$

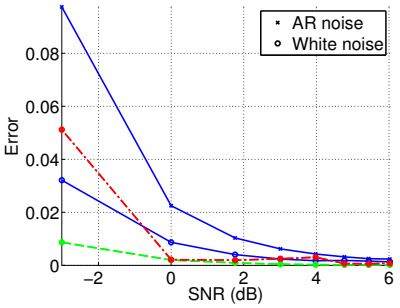
The smaller the error, the better the estimation of the lags and of the evoked potential.

Results with 3 different templates (cf. figure 7.14(a)) are presented in figure 7.14(b). Simulations were performed with two different types of noise (uncorrelated white noise and noise computed with an autoregressive filter whose coefficients were fitted on spontaneous EEG activity.

First it can be observed that in the noiseless case, the error is equal to 0. This validates the method and demonstrates that the estimation procedure is unbiased. Second it can be noticed that the best performance is obtained with uncorrelated white noise and that the errors remain small even with low SNR.



(a) The 3 different templates used in the simulations.



(b) Errors (see (7.7)) obtained with the 3 templates in (a) and 2 different types of noise (uncorrelated white noise and noise computed with an eighth-order autoregressive filter whose coefficients were fitted on spontaneous EEG activity).

Figure 7.14: Simulation results and errors estimates with different types of evoked responses ($\sigma_{lag} = 50$ ms). These values are the average errors obtained out of 10 repetitions of the experiment. It can be observed that errors are equal to 0 with no noise. The worst error 0.1 is observed with $SNR = -3dB$ on the template used in previous simulations. A correlation error below 0.1 can be considered as satisfactory.

7.6 DISCUSSION

By making use of advanced graph-based methods, we have proposed a robust and very fast two-step procedure for estimating the variability of evoked neural responses on single-trial MEG or EEG data: *spectral reordering* by Laplacian embedding followed by an estimation procedure by graph cuts. The whole process runs in a few seconds on real datasets of several hundred trials covering several hundred time points. The approach is a model-free, "data driven" algorithm, offering guarantees of global optimality for both of the steps. It does not suffer from initialization problems and does not assume a model, *e.g.,* imposing that the data can be well represented in an *a priori* dictionary of waveforms [16]. The procedure has several parameters that can be set automatically, as explained. Finally, numerical experiments on synthetic datasets confirm the robustness to noise of the full procedure.

This contribution puts an emphasis on the latency estimation problem. However, as illustrated in figure 7.5(c), the manifold learning method can handle other types of variability, for whose estimation a graph cut procedure could be designed. As long as the model of a 1D manifold holds, *i.e.,* that the variability can be parameterized by a single parameter, the methodology of this chapter can be applied.

Quantifying the variability of brain response delays, non invasively, in humans, can help to improve our understanding of the cognitive treatment of information. As presented in this chapter, once the delay has been corrected on each trial, the data can be realigned to the neural response, improving the quality of the estimated evoked response. A better control of the temporal aspects of the signal can moreover improve the spatial precision of the source reconstructions when solving the inverse problem [11].

Another application of single-trial estimation is related to the correlations between different channels, coming from the correlations between latencies of different functional brain regions. By computing delays of response independently on different channels, interactions between delays can be investigated. We can hypothesize that two neural processes that are independent exhibit uncorrelated delays while two sequential neural tasks have highly correlated delays. To quantify such interactions between various brain functional areas, simple correlations can be computed. It would be interesting to investigate what difference the delay correction, based on the data of one channel, makes to the amplitude of evoked potential on other channels. An augmentation of this amplitude on an other channel could be interpreted as a strong correlation between them, while an absence or very weak modification of this amplitude would suggest on the contrary an absence of correlation or a very weak one. Such questions, raised in the cognitive neurosciences community, can be addressed with the help of the method proposed in this chapter.

Although our method was applied to time series coming from a single EEG channel, it is also possible to estimate the delay of activations of an ICA component or more generally any source configuration. By employing signal-space projectors [208], or by projecting the full EEG or MEG recordings onto the forward field of a source configuration, we obtain a time series for each repetition of the experiment, which is what our algorithm requires as input.

The method detailed in this chapter has been implemented as an EEGLAB plugin. The code snippet in table 7.2 details how the plugin can be used from a Matlab script.

```
% Example Matlab code for single-trial latency estimation
% Load data
load('data/oddball3-num1-512Hz-chan10.set','-mat');

% Set parameters
use_ica = false; % Set to true, if you want to realign based on an ICA component
channel = 1; % Index of channel or ICA component used for realignment
time_win = [150 500]; % (ms) : work on this time window
bad_trials = []; % set bad trials

clear options
options.sigma = [0.01:0.01:0.2];
options.alpha = [0.001,0.01,0.1];
[EEG, com, order, lags, event_type, E_lags] = ...
    pop_extractlag( EEG , use_ica, channel, time_win, options);

% View ERP image reordered
figure;
pop_erpimage(EEG,1, [channel],[],EEG.chanlocs(channel).labels,1,1, ...
              { event_type }, [],'latency' ,'yerplabel','\muV','erp','cbar');

% Re-epoch the data
EEG = pop_epoch( EEG, {event_type}, [-0.4 0.3]);

% View ERP image of re-epoched data
figure;
pop_erpimage(EEG,1, [channel],[],EEG.chanlocs(channel).labels,1,1,{},[],'', ...
              'yerplabel','\muV','erp','cbar');
```

Table 7.2: Running the lag extraction pipeline on an EEGLAB dataset from the command line. Source code of the EEGLAB plug-in is available on the INRIA Forge `https://gforge.inria.fr/projects/eeglab-plugins/`.

7.7 Conclusion

The method presented in this chapter provides a computationally efficient and principled framework to address the challenging problem of estimating parameters on single-trial M/EEG data. Single-trial data analysis is an important goal in the M/EEG community as such analysis can give access to estimates that are not biased by the averaging process used with classical ERP studies.

The source code and the demo scripts necessary to reproduce the figures of this chapter are available in a Matlab EEGLAB Plug-in:

`https://gforge.inria.fr/projects/eeglab-plugins/`

Conclusion

In this thesis, the main methodological and theoretical aspects of M/EEG data processing have been covered, from accurate solution of the forward problem to efficient approaches to inverse problem in the context of distributed source models, including the challenging problem of single-trial data processing. We have been the main contributor to an open source software project, OpenMEEG, that offers to the M/EEG community the most accurate BEM solver available today.

Our work with experimental data led us to the analysis and implementation of state-of-the-art inverse solvers, domain to which we contributed by introducing a framework that enables to include multiple experimental conditions simultaneously. This work was motivated by the ambition to demonstrate that retinotopic mapping was possible with MEG. This topic was investigated from the design of an experimental protocol and the exploration of the data to the construction of principled methodology that allowed us to obtain promising results even if the final objective that consists in achieving timing of cortical processing in the visual cortex with only MEG is still for us an open problem.

Our interest on the research area motivated us to address some hard and still open questions in the field. Going beyond simple localization, we proposed a tracking algorithm working on triangular meshes that offers interesting perspectives for the investigation of cortical dynamics. We applied this method to visual processing and somatosensory MEG data which demonstrated that such an approach could provide insight into the timing of cortical processing.

The last topic addressed during this thesis concerns the problem of extracting information on single-trial M/EEG data, which is an issue of major interest, as such methods can give access to neural response estimates that are not biased by the averaging process used with classical ERP studies.

The contributions are thus threefold: theoretical, methodological and applied. Throughout this thesis, we tried to make the right mathematical choices to model the problems of interest. We believe this enabled us to propose appropriate and efficient algorithms so that we could finally tackle challenging neuroscience questions.

To summarize:

- We contributed to provide to the M/EEG community the most precise forward problem solver when considering realistic head models with piecewise constant conductivities.
- We presented the mathematical and computational details of the state-of-art inverse problem methods. Our implementation of all these methods is freely available in an open source project called EMBAL . The community has now access to simple but very efficient convex optimization schemes that we hope will contribute to the widespread use of such methods.
- We developed a framework for M/EEG inverse modeling able to integrate as *a priori* anatomo-functional knowledge between experimental conditions.
- We contributed to set up a full experimental study from protocol design and data exploration to the construction of a data analysis pipeline that offers promising results for the study of the visual cortex with MEG.
- We proposed a novel approach to address the hard problem of single-trial data analysis. We believe that this contribution can be a valuable tool to investigate inter-trial variabilities, which is of major interest for cognitive neuroscience studies.

Finally, we hope that this thesis elucidates some aspects of M/EEG data processing in order to improve the understanding but also the use of advanced methodological tools in the community. Consequently, we hope that such a better understanding will improve the quality

of results obtained with EEG and MEG in order for these brain functional imaging modalities to have a higher impact on both basic neuroscience and clinical studies.

Research Perspectives

Single-trial analysis

In this thesis, we approached the challenging problem of single-trial data analysis. This topic has a major interest especially in cognitive studies where the inter-trial variability can provide valuable information. In the near future, we plan to apply our existing tools for delay estimation to explore the human motor system. The next methodological step would be to extend our approach to various kinds of inter-trial variabilities. This work is currently starting in collaboration with Boris Burle in Marseille at CNRS / Université de Provence.

We would also be interested in single-trial inverse modeling with non-linear inverse solvers. When considering linear inverse solvers, averaging the single-trial estimates or inverting the averaged M/EEG measurements provides the same result. With non linear inverse solvers, this does not hold. We would like to approach the problem using sparsity inducing priors where the penalization would be on the time-frequency decompositions. M/EEG signals are oscillatory. Therefore, using a time-frequency representation to constrain the inverse solvers seems very reasonable.

Another problem we would like to address is “resting state M/EEG”. While in the last two research problems the recordings were time locked to the beginning of the stimulation, resting state M/EEG would impose to work on raw data without “time triggers” such as the stimulus onsets. Independent component analysis (ICA) is an approach to address this problem but it suffers from some practical problems like the necessity to set *a priori* the number of components. The approach we plan to consider is based on a research topic generally referred to as “sparse coding”. Our preliminary results on this topic seem to confirm the power of the approach.

Investigations of the visual system with MEG

The neuroscience topic that motivated this thesis is the understanding of human visual system using M/EEG. To go one step further from where we arrived on the project of retinotopic mapping with MEG, we would like to conduct more experiments and investigate the role of the different stimulation parameters on the activation maps obtained. We would like to investigate the comparison of mapping results using either classical visual evoked potentials or steady state stimulations as in chapter 5. This would challenge the robustness and ultimately improve our data processing pipeline for retinotopic mapping with MEG. If our pipeline reaches a limit, we might consider going towards a generative model of the structure of V1 and V2 to better constrain the inverse problem. To conclude, we believe that a necessary condition for such a project to succeed is to have more interaction with experimentalists in order to benefit from their expertise to achieve a good experimental control and an easy low level processing of the data.

With well-designed experiments and robust data processing pipelines, the challenging objective that is to achieve precise estimations of cortical dynamics could be addressed. Potential approaches to this problem are based the analysis of phase differences as mentioned in chapter 5 or using tracking methods like in chapter 6.

Multi-conditions inverse modeling

During this thesis, we investigated the use of multiple experimental conditions simultaneously within the inverse modeling. The classical way consists in applying inverse solver to the data from each condition individually and in a second step comparing the source estimates obtained for each of them. However, this approach has limitations especially when considering inverse modeling with sparse priors.

Being able to compare experimental conditions is of major interest for brain research. A cognitive question that could be answered is: Is this part of the brain activated by condition 1 and condition 2 simultaneously? Or equivalently what activation pattern is shared between condition 1 and 2, and what is different in the cognitive process?

We think our expertise on the field of M/EEG inverse modeling provides us good tools to address such questions in a mathematically principled and computationally efficient way.

Improving the M/EEG data processing pipelines

As it appears clearly in this thesis, we devoted a lot of time to develop and disseminate software, to detail implementation problems, and to provide practical tips to help analyze M/EEG data. The motivation of this comes from the observation that analyzing M/EEG data is difficult, time consuming, and sometimes even despairing when the results are not as nice as expected.

When confronted with real data, we want to be confident about the tools we use. We do not want to permanently wonder if the low quality of the result is due to the bad quality of the data, to a bad choice of a method or, worse, to a "buggy" implementation.

To avoid this, the challenge is to set up M/EEG data processing pipelines whose building blocks are reliable and easy to use. These building blocks start from data processing and removal of artefacts, to the accurate and automatic forward modeling with realistic head models, but also the fast and efficient computation of inverse solvers.

The best examples today of freely available software packages that attempt to achieve these goals are EEGLAB, Fieldtrip, Brainstorm, and MNE. Each of these packages have different ambitions, a different usability for a non expert, a different level of automatization, and a different level of flexibility when comes the problem of integrating new tools in the processing pipelines.

We believe that M/EEG research will benefit from the improvement of these processing pipelines. By sharing data and software and by standardizing pipelines, one could guarantee the reproducibility of the results and facilitate the comparison between methods. Sometimes, we wonder if M/EEG research is not as popular as fMRI because of the complexity of each of the steps to go from raw MEG data to clean cortical activations. The different steps require mathematical and computing skills, a relatively good understanding of the physics, and an ability to interpret the results for neuroscience. And as it is known "A chain breaks at its weakest link".

Appendix

APPENDIX A

KRONECKER PRODUCTS

Kronecker product is a very convenient tool that often enables compact and readable matrix computations.

Definition A.1. Kronecker product Let $\mathbf{A} \in \mathbb{R}^{m \times n}$ and $\mathbf{B} \in \mathbb{R}^{p \times q}$. Then the Kronecker product (or tensor product) of $\mathbf{A}$ and $\mathbf{B}$ is defined as the matrix

$$\mathbf{A} \otimes \mathbf{B} = \begin{pmatrix} a_{11}\mathbf{B} & \cdots & a_{1n}\mathbf{B} \\ \vdots & \ddots & \vdots \\ a_{m1}\mathbf{B} & \cdots & a_{mn}\mathbf{B} \end{pmatrix} \in \mathbb{R}^{mp \times nq} \tag{A.1}$$

PROPERTIES OF KRONECKER PRODUCTS

Theorem A.1. *Let* $\mathbf{A} \in \mathbb{R}^{m \times n}, \mathbf{B} \in \mathbb{R}^{r \times s}, \mathbf{C} \in \mathbb{R}^{n \times p}$ *and* $\mathbf{D} \in \mathbb{R}^{s \times t}$. *Then*

$$(\mathbf{A} \otimes \mathbf{B})(\mathbf{C} \otimes \mathbf{D}) = (\mathbf{AC} \otimes \mathbf{BD}) \in \mathbb{R}^{mr \times pt} \ .$$

Theorem A.2. *For all* $\mathbf{A}$ *and* $\mathbf{B}$, $(\mathbf{A} \otimes \mathbf{B})^T = \mathbf{A}^T \otimes \mathbf{B}^T$

Theorem A.3. *If* $\mathbf{A}$ *and* $\mathbf{B}$ *are non singular,* $(\mathbf{A} \otimes \mathbf{B})^{-1} = \mathbf{A}^{-1} \otimes \mathbf{B}^{-1}$

Theorem A.4. *Let* $\mathbf{A} \in \mathbb{R}^{m \times n}$ *have a singular value decomposition* $\mathbf{U}_A \mathbf{\Sigma}_A \mathbf{V}_A^T$ *and let* $\mathbf{B} \in \mathbb{R}^{p \times q}$ *have a singular value decomposition* $\mathbf{U}_B \mathbf{\Sigma}_B \mathbf{V}_B^T$. *Then*

$$(\mathbf{U}_A \otimes \mathbf{U}_B)(\mathbf{\Sigma}_A \otimes \mathbf{\Sigma}_B)(\mathbf{V}_A^T \otimes \mathbf{V}_B^T)$$

yields a singular value decomposition of $\mathbf{A} \otimes \mathbf{B}$ *(after a simple reordering of the diagonal elements of* $\mathbf{\Sigma}_A \otimes \mathbf{\Sigma}_B$ *and the corresponding right and left singular vectors).*

Let $\mathbf{A} \in \mathbb{R}^{m \times n}$. The matrix $\mathbf{A}$ can be converted to a vector by *stacking* all columns of $\mathbf{A}$ on top of one another. Let $\mathbf{a}_{\cdot i}$ denote the ith column of $\mathbf{A}$.

$$\text{vec}(\mathbf{A}) = \begin{pmatrix} \mathbf{a}_{\cdot 1} \\ \vdots \\ \mathbf{a}_{\cdot n} \end{pmatrix} \in \mathbb{R}^{mn}$$

Proposition A.5. *Let* $\mathbf{A} \in \mathbb{R}^{m \times n}$ *and* $\mathbf{B} \in \mathbb{R}^{p \times q}$ *and* $\mathbf{X} \in \mathbb{R}^{n \times p}$. *Then*

$$vec(\mathbf{AXB}) = (\mathbf{B}^T \otimes \mathbf{A})vec(\mathbf{X}) \ .$$

The proposition provides the following result, allowing to compute the product of a vector with a Kronecker product without actually assembling the full matrix $\mathbf{A} \otimes \mathbf{B}$.

For $\mathbf{x} \in \mathbb{R}^{nq}$,

$$(\mathbf{A} \otimes \mathbf{B})\mathbf{x} = \text{vec}(\mathbf{B}\text{mat}(\mathbf{x})\mathbf{A}^T)$$

where the notation $\text{mat}(x)$ denotes the matrix in $\mathbb{R}^{q \times n}$ such that $\text{vec}(\text{mat}(x))$ is equal to x.

APPENDIX B

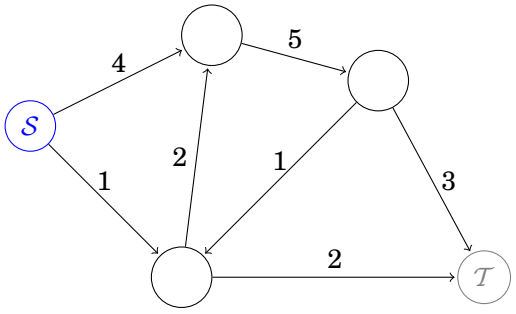
INTRODUCTION TO GRAPH-CUTS

Lets consider an oriented graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V}$ is the set of vertices, often called nodes, and $\mathcal{E} \subset \mathcal{V}^2$ is the set of oriented edges, *i.e.,* $(a, b) \neq (b, a)$.

Lets consider the function $w : \mathcal{E} \to \mathbf{R}^+ \cup +\infty$, that assigns a weight, also called capacity, to each edge. Notice that values of w are necessarily positive.

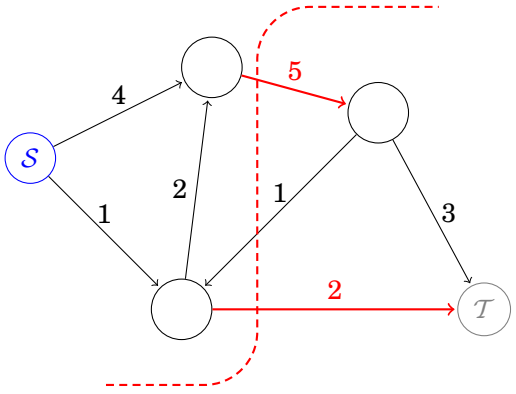
Among $\mathcal{E}$ are two particular vertices, $\mathcal{S}$ and $\mathcal{T}$. $\mathcal{S}$ is called the *source* and does not have any incoming edges while $\mathcal{T}$ is called the *sink* and does not have any outgoing edges.

Here is an example of such a graph:



Definition B.1 (Cut)**.** : A cut $(\mathbf{S}, \mathbf{T})$ of the graph $\mathcal{G}$ is a partition of the vertices (*i.e.,* $\mathbf{S} \cup \mathbf{T} = \mathcal{V}$ and $\mathbf{S} \cap \mathbf{T} = \emptyset$) such that $\mathcal{S} \in \mathbf{S}$ and $\mathcal{T} \in \mathbf{T}$

Here is an example of cut:



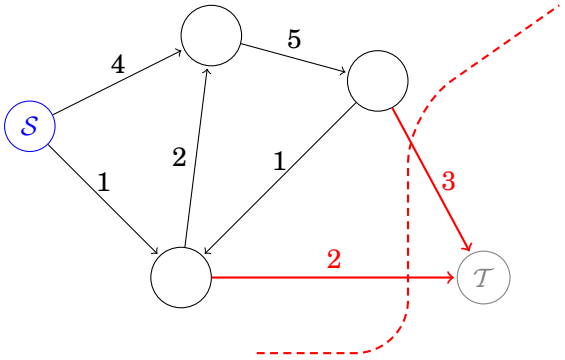
Definition B.2 (Weight of a cut)**.** : The weight of a cut $(\mathbf{S}, \mathbf{T})$ is defined as

$$c(\mathbf{S}, \mathbf{T}) = \sum_{\substack{(p,q)\in\mathcal{E} \\ p\in\mathbf{S}, q\in\mathbf{T}}} w(p, q) \tag{B.1}$$

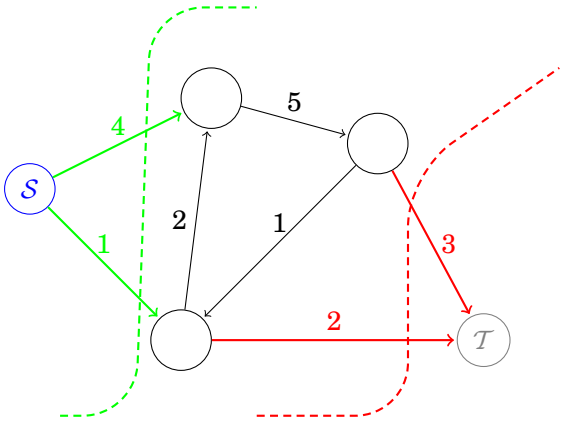
The weight of the cut presented in the figure above is given by the sum of the weights of the red colored edges, *i.e.,* $2 + 5 = 7$.

Definition B.3 (Minimum cut - *MinCut*)**.** : A cut is minimal if the weight of the cut is not larger than the weight of any other cut.

The following figure presents a *MinCut* of G. Its weight is $2 + 3 = 5$.



The minimum cut might however not be unique. In the figure below the two represented cuts are both minimum with a weight of 5.



One of the fundamental results in combinatorial optimization is that the minimum cut problem can be solved by finding a maximum flow from the source $\mathcal{S}$ to the sink $\mathcal{T}$. Speaking informally, maximum flow is the maximum "amount of water" that can be sent from the source to the sink by interpreting graph edges as directed "pipes" with capacities equal to edge weights. The theorem of Ford and Fulkerson [53] states that a maximum flow from $\mathcal{S}$ to $\mathcal{T}$ saturates a set of edges in the graph dividing the nodes into two disjoint parts, S and T, corresponding to a minimum cut. Thus, *MinCut* and *MaxFlow* problems are equivalent. In fact, the maximum flow value is equal to the cost of the minimum cut.

Presented more formally it leads to:

Definition B.4 (Flow)**.** Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be a graph, w its capacity function, $\mathcal{S}$ and $\mathcal{T}$ the source and the sink. A *flow* is a function $f : \mathcal{E}^* \rightarrow \mathbf{R}_+$ ($\mathcal{E}^*$ is the set of edges and their inverse) satisfying the following properties:

- for each edge $e = (p, q) \in \mathcal{E}$

$$f(p, q) = -f(q, p) \tag{B.2}$$

- for each vertex p besides $\mathcal{S}$ and $\mathcal{T}$,

$$\sum_{\substack{e=(p,.) \\ e \in \mathcal{E}^*}} f(e) = 0 \tag{B.3}$$

- for each edge $e \in \mathcal{E}$,

$$f(e) \leq w(e) \tag{B.4}$$

The constraint in equation (B.3) corresponds to a conservation law similar to the Kirchoff law. The constraint in equation (B.4) imposes the flow in edge e to be smaller then its capacity $w(e)$. Both equation (B.3) and equation (B.4) imply that:

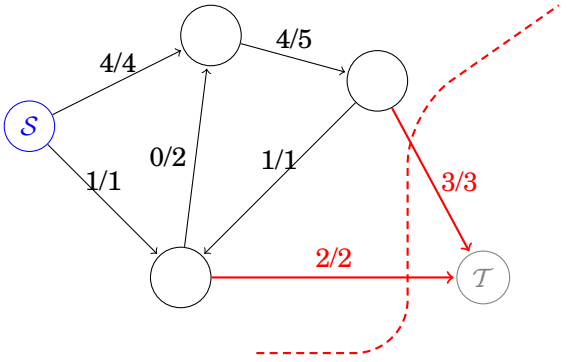
$$\sum_{e=(\mathcal{S},.)} f(e) = \sum_{e=(.,\mathcal{T})} f(e) \tag{B.5}$$

Equivalently, the amount of liquid that comes out of the source $\mathcal{S}$ is equal to the amount of liquid that goes into the sink $\mathcal{T}$. The quantity is called the value of the flow.

Definition B.5 (Maximum flow - *MaxFlow*)**.** : A flow is maximum if its value is not smaller than the value of any other flow.

Theorem B.1 (*MinCut* - *MaxFlow* equivalence)**.** *The* MinCut *of a graph $\mathcal{G}$ as defined above, is equal to the* MaxFlow *[53].*

A *MaxFlow* on the example graph with a corresponding *MinCut*:



Our interest for problems that can be reformulated as a minimum cut problem comes from the following theorem.

Theorem B.2 (*MinCut* - *MaxFlow* complexity)**.** *Finding the maximum flow, and equivalently the minimum cut, of a graph is a problem that can be solved in polynomial time.*

In other words, this theorem implies that the *MinCut* problems are “efficiently solvable” or “tractable”. In practice, minimum cuts are not are obtained via computation of maximum flows.

Algorithms for the *MinCut* and *MaxFlow* Problem

There are many standard polynomial time algorithms for *MinCut*/*MaxFlow* [40]. These algorithms can be divided into two main groups: “push-relabel” style methods [89] and algorithms based on augmenting paths. In practice the push-relabel algorithms perform better for general graphs. In vision applications, however, the most common type of a graph is a two or a higher dimensional grid. For regular graphs like grids, Boykov and Kolmogorov [22] developed a fast augmenting path algorithm which often significantly outperforms the push-relabel algorithm. Furthermore, its observed running time is linear.

We now explain briefly how the augmenting path algorithm works. Given a flow f, the residual capacity $r(p,q)$ of an edge $e = (p,q) \in \mathcal{E}$ linking node p to node q is the maximum additional flow that can be sent from node p to node q using the edges (p,q) and (q,p). The residual capacity $r(p,q)$ has two components: the unused capacity of the edge (p,q): $w(e) - f(e)$

and the current flow $f(q,p)$ from node q to node p which can be reduced to increase the flow from p to q. A residual graph $\mathcal{G}(f)$ of a graph $\mathcal{G}$ consists of the node set $\mathcal{V}$ and the edges with positive residual capacity (with respect to the flow f). The topology of $\mathcal{G}(f)$ is identical to $\mathcal{G}$. $\mathcal{G}(f)$ differs only in the capacity of its edges and so for zero flow, *i.e.,* $f(p,q) = 0 \, \forall (p,q) \in \mathcal{E}$, $\mathcal{G}(f)$ is same as $\mathcal{G}$. An augmenting path is a path from the source to the sink along unsaturated edges of the residual graph. Augmenting path based algorithms for solving the max-flow problem work by repeatedly finding augmenting paths in the residual graph and saturating them. When no more augmenting paths can be found, *i.e.,* the source and sink are disconnected in the residual graph, the maximum flow is obtained.

APPENDIX C

TIME FREQUENCY ANALYSIS WITH GABOR FILTERS

Gabor filters are linear filters localized in time and frequency. In time, they consist of complex exponential functions modulated by a Gaussian with standard deviation σ. The parameter σ controls the trade-off between temporal precision and spectral precision of the filter.

Let $\psi^{\sigma}_{t_0,f_0}$ denote the Gabor filter centered at time t_0 and at frequency f_0 (cf. figure C.1) [143].

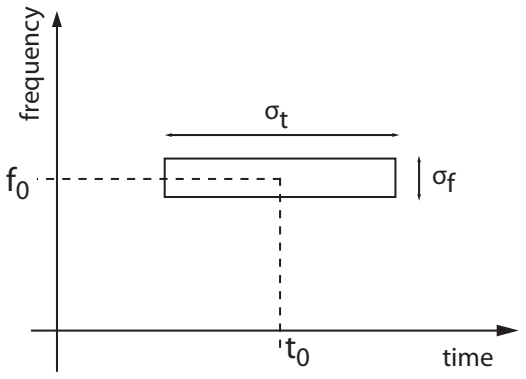


Figure C.1: Spectral support of the Gabor filter $\psi^{\sigma}_{t_0,f_0}$. The parameter σ in the text corresponds to σ_t in the figure.

The expression of $\psi^{\sigma}_{t_0,f_0}$ is given by:

$$\psi^{\sigma}_{t_0,f_0}(t) = (\pi\sigma^2)^{-1/4} e^{2i\pi f_0(t-t_0)} e^{-\frac{(t-t_0)^2}{2\sigma^2}} \quad .$$

and its Fourier transform is given by:

$$\hat{\psi}^{\sigma}_{t_0,f_0}(f) = (4\pi\sigma^2)^{-1/4} e^{-2i\pi f t_0} e^{-\frac{\sigma^2}{2}(2\pi(f-f_0))^2} \quad .$$

The temporal resolution of $\psi^{\sigma}_{t_0,f_0}$ is denoted σ and by application of the Fourier transform, one can observe that its spectral resolution is proportional to $1/\sigma$. It means that the area of the box in figure C.1 is constant whatever the choice of σ. In order to be precise in time one needs to reduce σ which implies to lose in spectral resolution.

M/EEG signals are oscillatory and typically consist of bursts of activations with a few oscillations. These oscillations can be observed on the raw signal, especially at low frequencies. For this reason, we prefer to parameterize Gabor filters with an oscillatory parameter ξ

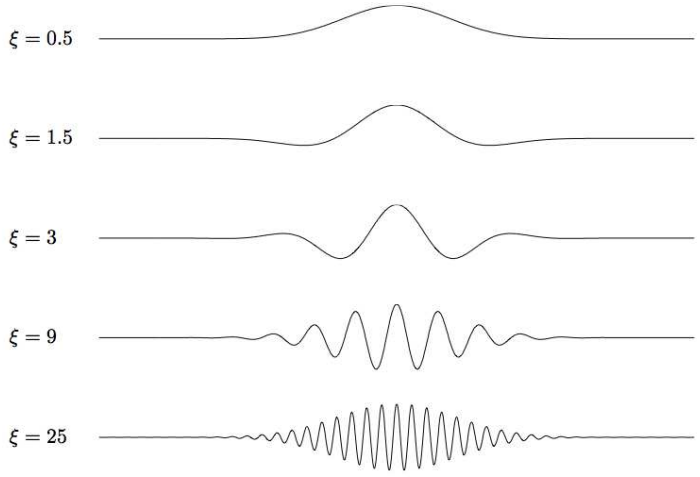


Figure C.2: Gabor atoms for different values of the oscillation parameter ξ (modified by varying f_0 with a contant σ). A low oscillation parameter produces a transient wave, and a high value a sustained oscillation.

rather than with σ [16]. The parameterization with σ is more classical in the signal processing community. The parameter ξ is defined by:

$$\xi = 2\pi f_0 \sigma \ .$$

The parameter σ stretches or compresses the time support of the filter without modifying its frequency, whereas ξ can be related to the number of visible oscillations of the filter (cf. figure C.2). When f_0 increases, the parameter σ decreases in order to maintain the number of oscillations ξ constant.

One can observe in figure C.3 an example of time frequency decomposition obtained with Gabor filters ($\xi = 10$). One can notice that the temporal resolution of the atoms increases with the frequency.

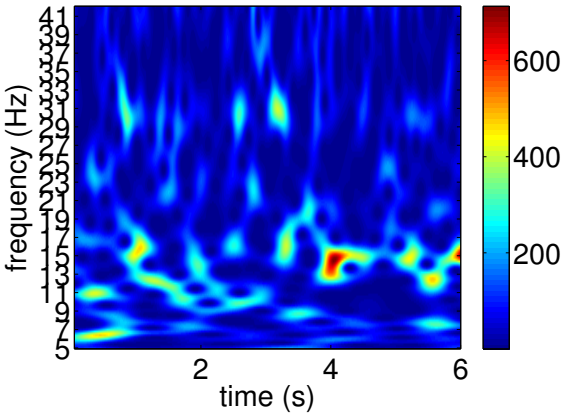


Figure C.3: Sample time frequency map, a.k.a. spectrogram, estimated with Gabor filters with $\xi = 10$ on real MEG data extracted from the retinotopy study (cf. chapter 5).

APPENDIX D

PUBLICATIONS OF THE AUTHOR

JOURNAL PAPERS

A. Gramfort, T. Papadopoulo, S. Baillet and M. Clerc, *Tracking cortical activity with spatio-temporal constraints using graph cuts*, SIAM Imaging Science, (submitted).

A. Gramfort, R. Keriven and M. Clerc, *Graph-based estimation of 1-D variability in event related neural responses*, IEEE Transactions on Biomedical Engineering (TBME), (submitted).

A. Gramfort and M. Kowalski, *M/EEG inverse problem with structured sparse priors: why and how.*, In preparation.

A. Gramfort, T. Papadopoulo, E. Olivi and M. Clerc, *OpenMEEG: opensource software for quasistatic bioelectromagnetics.*, In preparation.

PEER-REVIEWED CONFERENCE PAPERS AND ABSTRACTS

A. Gramfort and M. Kowalski, *Improving M/EEG source localization with an inter-condition sparse prior*, Proceedings International Symposium on Biomedical Imaging: From Nano to Macro (ISBI), jun. 2009.

M. Kowalski and A. Gramfort, *A priori par normes mixtes pour les problèmes inverses: Application à la localisation de sources en M/EEG*, Proceedings GRETSI, sept. 2009.

B. Cottereau, J. Lorenceau, A. Gramfort, M. Clerc, B. Thirion and S. Baillet, *Fine chronometric mapping of human visual areas*, Human Brain Mapping, jun. 2009.

A. Gramfort, T. Papadopoulo, B. Cottereau, S. Baillet and M. Clerc, *Tracking cortical activity with spatio-temporal constraints using graph cuts*, International Conference on Biomagnetism (BIOMAG), aug. 2008.

B. Cottereau, A. Gramfort, J. Lorenceau, M. Clerc, B. Thirion and S. Baillet, *Fast retinotopic mapping of visual fields using MEG*, Human Brain Mapping, jun. 2008.

A. Gramfort, B. Cottereau, M. Clerc, B. Thirion and S. Baillet, *Challenging the estimation of cortical activity from MEG with simulated fMRI-constrained retinotopic maps*, Proceedings of

the 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 4945-4948, aug 2007.

A. Gramfort and M. Clerc, *Low dimensional representations of MEG/EEG data using Laplacian Eigenmaps*, Proceedings Noninvasive Functional Source Imaging of the Brain and Heart (NFSI), 169-172, oct. 2007.

M. Clerc, A. Gramfort, P. Landreau and T. Papadopoulo, *MEG and EEG processing with OpenMEEG*, Proceedings of Neuromath, 2007.

B. Cottereau, J. Lorenceau, A. Gramfort, B. Thirion, M. Clerc and S. Baillet, *Fast Retinotopic Mapping of Visual Fields using MEG*, Proceedings of Neuromath, 2007.

SOFTWARE

OpenMEEG: C++ package to solve the M/EEG forward problem with the symmetric boundary element method.

`https://gforge.inria.fr/projects/openmeeg`

EMBAL: Matlab Toolbox for M/EEG inverse modeling with distributed source models.

`https://gforge.inria.fr/projects/embal`

Matlab EEGLAB Plug-in for single-trial parameter estimation.

`https://gforge.inria.fr/projects/eeglab-plugins/`

Bibliography

[1] Beck A and Teboulle M. Fast iterative shrinkage-thresholding algorithm for linear inverse problems. *SIAM J. Imaging Sciences*, 2:183 – 202, 2009.

[2] Gramfort A., Cottereau B., Clerc M., Thirion B., and Baillet S. Challenging the estimation of cortical activity from MEG with simulated fMRI-constrained retinotopic maps. In *EMBC 2007: Proceedings of the 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Jun 2007.

[3] J P R Bolton A A Ioannides and C J S Clarke. Continuous probabilistic solutions to the biomagnetic inverse problem. *Inverse Problems*, 6:523–542, 1990.

[4] G Adde. *Méthodes de traitement d'image appliqués au problème inverse en Magnéto-Electro-Encéphalographie*. PhD thesis, Ecole Nationale des Ponts et Chaussées, 2005.

[5] G Adde, M Clerc, and R Keriven. Imaging methods for MEG/EEG inverse problem. In *Proc. Joint Meeting of 5th International Conference on Bioelectromagnetism and 5th International Symposium on Noninvasive Functional Source Imaging*, 2005.

[6] Cottereau B. *Modèles hiérarchiques en imagerie MEG/EEG - Application à la création rapide de cartes rétinotopiques*. PhD thesis, Université Paris-Sud 11, May 2008.

[7] B. Thirion and O. Faugeras. Nonlinear dimension reduction of fMRI data: the Laplacian embedding approach. In *Proceedings ISBI*, pages 372–375, Apr 2004.

[8] B. Thirion, S. Dodel, and J.-B. Poline. Detection of signal synchronizations in resting-state fMRI datasets. *NeuroImage*, 29:321–327, Aug 2005.

[9] S Baillet and L Garnero. A bayesian approach to introducing anatomo-functional priors in the eeg/meg inverse problem. *Biomedical Engineering*, 44(5), Jan 1997.

[10] S. Baillet, J. C. Masher, and R. M. Leahy. Electromagnetic brain imaging using BrainStorm. In *Biomedical Imaging: Nano to Macro, 2004. IEEE International Symposium on*, volume 1, pages 652–655, 2004.

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

[12] J. Bect, L. Blanc-Féraud, G. Aubert, and A. Chambolle. A l1-unified variational framework for image restoration. In T. Pajdla and J. Matas, editors, *Proc. European Conference on Computer Vision (ECCV)*, volume LNCS 3024, pages 1–13, Prague, Czech Republic, May 2004. Springer.

[13] Murat Belge, Misha E. Kilmer, and Eric L. Miller. Efficient Determination of Multiple Regularization Parameters in a Generalized L-Curve Framework. *Inverse Problems*, 18:2002, 2002.

[14] M. Belkin and P. Niyogi. Laplacian eigenmaps for dimensionality reduction and data representation. *Neural Computation*, 15(6):1373–1396, jun 2003.

[15] C. Bénar, M. Clerc, and T. Papadopoulo. Adaptive time-frequency models for single-trial M/EEG analysis. In Karssemeijer and Lelieveldt, editors, *Information Processing in Medical Imaging*, volume 4584 of *Lecture Notes in Computer Science*, pages 458–469. Springer, 2007.

[16] C. Bénar, T. Papadopoulo, B. Torrésani, and M. Clerc. Consensus matching pursuit for multi-trial EEG signals. *Journal of Neuroscience Methods*, 180(1):161–170, 2009.

[17] C.G. Bénar, D. Schön, S. Grimault, B. Nazarian, B. Burle, M. Roth, J.M. Badier, P. Marquis, C. Liégeois-Chauvel, and J.L. Anton. Single-trial analysis of oddball event-related potentials in simultaneous EEG-fMRI. *Human Brain Mapping*, 28:602–613, 2007.

[18] P. Berg and M. Scherg. A fast method for forward computation of multiple-shell spherical head models. *Electroencephalogr. Clin. Neurophysiol.*, 90(1):58–64, 1994.

[19] D.A. Boas, D.H. Brooks, E.L. Miller, C.A. DiMarzio, M. Kilmer, R.J. Gaudette, and Quan Zhang. Imaging the body with diffuse optical tomography. *Signal Processing Magazine, IEEE*, 18(6):57–75, Nov 2001.

[20] Stephen Boyd and Lieven Vandenberghe. *Convex Optimization*. Cambridge University Press, March 2004.

[21] Y Boykov and M Jolly. Interactive graph cuts for optimal boundary and region segmentation of objects in ndimages. *International Conference on Computer Vision*, 1:115, Jan 2001.

[22] Y Boykov and V Kolmogorov. An Experimental Comparison of Min-Cut/Max-Flow Algorithms for Energy Minimization in Vision. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(9), Sep 2004.

[23] Y. Boykov and O. Veksler. *Mathematical Models in Computer Vision: The Handbook. N. Paragios, Y. Chen and O. Faugeras Eds.*, chapter Graph Cuts in Vision and Graphics: Theories and Applications. Springer, 2006.

[24] A.A Brewer, J.L., A R Wade, and B A Wandell. Visual field maps and stimulus selectivity in human ventral occipital cortex. *Nature Neuroscience*, 8(8):1102–1109, 2005.

[25] K. Brodmann. *Vergleichende Lokalisationslehre der Grobhirnrinde*. J.A.Barth, Leipzig, 1909.

[26] DH Brooks, GF Ahmad, RS MacLeod, and GM Maratos. Inverse electrocardiography by simultaneous imposition of multiple constraints. *IEEE transactions on biomedical engineering*, 46(1):3–18, 1999.

[27] A Bruce, S Sardy, and P Tseng. Block coordinate relaxation methods for nonparamatric signal denoising. *Proceedings of SPIE*, 3391(75), Jan 1998.

[28] J Bullier. Integrated model of visual processing. *Brain Res. Reviews*, 36:96–107, 2001.

[29] L Chalupa and J.S Werner. *The visual neurosciences*. The MIT Press, 2004.

[30] A Chambolle. An algorithm for total variation minimization and applications. *Journal of Mathematical Imaging and Vision*, 20(1-2):89–97, Jan 2004.

[31] A. Chambolle and P. L. Lions. Image recovery via total variation minimization and related problems. *Numer. Math.*, 76:167–188, 1997.

[32] Tony F. Chan, Gene H. Golub, and Pep Mulet. A nonlinear primal-dual method for total variation-based image restoration. *SIAM J. Sci. Comput.*, 20(6):1964–1977, 1999.

[33] N. Chauveau, X. Franceries, B. Doyon, B. Rigaud, J.P. Morucci, and P. Celsis. Effects of skull thickness, anisotropy, and inhomogeneity on forward EEG/ERP computations using a spherical three-dimensional resistor mesh model. *Human Brain Mapping*, 21:86–97, 2004.

[34] S Chen, D Donoho, and M Saunders. Atomic decomposition by basis pursuit. *SIAM Journal on Scientific Computing*, Jan 1999.

[35] David Cohen. Magnetoencephalography: Evidence of magnetic fields produced by alpha-rhythm currents. *Science*, 161(3843):784–786, August 1968.

[36] David Cohen. Magnetoencephalography: Detection of the brain's electrical activity with a superconducting magnetometer. *Science*, 175(4022):664–666, February 1972.

[37] R.R Coifman, S Lafon, A.B Lee, M Maggioni, and Nadler. Geometric diffusions as a tool for harmonic analysis and structure definition of data: Diffusion maps. *Proceedings of the National Academy of Sciences*, 102(21):7426–7431, 2005.

[38] Y. Cointepas, J.-F. Mangin, Line Garnero, J.-B. Poline, and H. Benali. BrainVISA: Software platform for visualization and analysis of multi-modality brain data. In *Proc. 7th HBM*, page S98, Brighton, United Kingdom, 2001.

[39] P. L. Combettes and V. R. Wajs. Signal recovery by proximal forward-backward splitting. *Multiscale Modeling and Simulation*, 4(4):1168–1200, November 2005.

[40] William J. Cook, William H. Cunningham, William R. Pulleyblank, and Alexander Schrijver. *Combinatorial Optimization.* John Wiley & Sons, 1998.

[41] Diego Cosmelli, Olivier David, Jean-Philippe Lachaux, Jacques Martinerie, Line Garnero, Bernard Renault, and Francisco Varela. Waves of consciousness: ongoing cortical patterns during binocular rivalry. *NeuroImage*, 23(1):128–140, September 2004.

[42] B. Cottereau, A. Gramfort, J. Lorenceau, B. Thirion, M. Clerc, and S. Baillet. Fast retinotopic mapping of visual fields using MEG. In *Human Brain Mapping*, 2008.

[43] B Cottereau, K Jerbi, and S Baillet. Multiresolution imaging of meg cortical sources using an explicit piecewise model. *Neuroimage*, Sep 2007.

[44] B. Cottereau, J. Lorenceau, A. Gramfort, M. Clerc, and S. Baillet. Fine chronometric mapping of human visual areas. In *Human Brain Mapping*, jun 2009.

[45] B. Cottereau, J. Lorenceau, A. Gramfort, B. Thirion, M. Clerc, and S. Baillet. Fast retinotopic mapping of visual fields using meg. In *Proceedings of Neuromath*, 2007.

[46] Benoit Cottereau, Jean Lorenceau, Alexandre Gramfort, Bertrand Thirion, Maureen Clerc, and Sylvain Baillet. Fast retinotopic mapping of visual fields using meg. In *Proceedings of Neuromath*, 2007.

[47] B.N. Cuffin. EEG localization accuracy improvements using realistically shaped head models. *IEEE Trans. on Biomed. Engin.*, 43(3), 1996.

[48] F.H Lopes da Silva, A van Rotterdam, P Barts, E van Heusden, and W Burr. Model of neuronal populations. the basic mechanism of rhythmicity. *M.A. Corner, D.F. Swaab (eds) Progress in brain research*, 45:281–308, 1976.

[49] A Dale, A Liu, B Fischl, and R Buckner. Dynamic statistical parametric neurotechnique mapping: combining fMRI and MEG for high-resolution imaging of cortical activity. *Neuron*, 26:55–67, 2000.

[50] A Dale and M Sereno. Improved localization of cortical activity by combining EEG and MEG with MRI cortical surface reconstruction. *Journal of Cognitive Neuroscience*, Jan 1993.

[51] Anders Dale, Martin Sereno, Bruce Fischl, Sean Marrett, Arthur Liu, Eric Halgren, Kevin Teich, Christian Haselgrove, Doug Greve, and Florent Segonne. Freesurfer manual.

[52] P.M Daniel and D Whitteridge. The representation of the visual field on the cerebral cortex in monkeys. *Journal of Neurophysiology*, 159:203–221, 1961.

[53] G. B. Dantzig and D. R. Fulkerson. On the max-flow min-cut theorem of networks. *Ann. Math. Studies*, 38, 1956.

[54] G. Dassios and F. Kariotou. Magnetoencephalography in ellipsoidal geometry. *Journal of Mathematical Physics*, 44:220–241, 2003.

[55] I Daubechies, M Defrise, and C De Mol. An iterative thresholding algorithm for linear inverse problems with a sparsity constraint. *Communications on Pure and Applied Mathematics*, Jan 2004.

[56] I Daubechies, R DeVore, M Fornasier, and S Gunturk. Iteratively re-weighted least squares minimization: Proof of faster than linear rate for sparse recovery. *Information Sciences and Systems*, 2008.

[57] J de Munck. A linear discretization of the volume conductor boundary integral equation using analytically integrated elements. *IEEE Trans. Biomed. Eng.*, 39(9):986–990, 1992.

[58] J.C. de Munck and M.J. Peters. A fast method to compute the potential in the multisphere model. *IEEE Trans. on Biomed. Engin.*, 40(11):1163–1174, 1993.

[59] Arnaud Delorme and Scott Makeig. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, 134(1):9–21, 2004.

[60] A. P. Dempster, N. M. Laird, and D. B. Rubin. Maximum likelihood from incomplete data via the em algorithm. *Journal of the Royal Statistical Society. Series B (Methodological)*, 39(1):1–38, 1977.

[61] D Donoho. De-noising by soft-thresholding. *IEEE Trans. Information Theory*, 41(3):613–627, May 1995.

[62] R Dougherty, V Koch, A Brewer, B Fischer, J Modersitzki, and B Wandell. Visual field representations and locations of visual areas v1/2/3 in human visual cortex. *Journal of Vision*, 3:586–598, 2003.

[63] Bradley Efron, Trevor Hastie, Lain Johnstone, and Robert Tibshirani. Least angle regression. *Annals of Statistics*, 32:407–499, 2004.

[64] S Engel, D Rumelhart, B Wandell, A Lee, G Glober, E-J Chichilnisky, and M Shadlen. fmri of human visual cortex. *Nature*, 369:525–529, 1994.

[65] S.A Engel, G.H Glover, and B.A Wandell. Retinotopic organization in human visual cortex and the spatial precision of functional mri. *Cerebral Cortex*, 7:181–192, 1997.

[66] D Van Essen, H Drury, S Joshi, and M Miller. Functional and structural mapping of human cerebral cortex: Solutions are in the surfaces. *Proceedings of the National Academy of Sciences*, 95:788–795, 1998.

[67] Sharbrough F, Chatrian G-E, Lesser RP, Lüders H, Nuwer M, and Picton TW. American Electroencephalographic Society Guidelines for Standard Electrode Position Nomenclature. *Journal of Clinical Neurophysiology*, 8:200–202, 1991.

[68] O Faugeras, F Clément, R Deriche, R Keriven, T Papadopoulo, J Roberts, T Viéville, F Devernay, J Gomes, G Hermosillo, P Kornprobst, and D Lingrand. The inverse EEG and MEG problems: The adjoint space approach I: The continuous case. Technical Report 3673, INRIA, 1999.

[69] I Fawcett, G Barnes, A Hillebrand, and K Singh. The temporal frequency tuning of human visual cortex investigated using synthetic aperture magnetometry. *Neuroimage*, Jan 2004.

[70] D.J Felleman and D.C Essen. Distributed hierarchical processing in the primate cerebral cortex. *Cereb Cortex*, 1:1–47, 1991.

[71] A Ferguson, X Zhang, and G Stroink. A complete linear discretization for calculating the magnetic field using the boundary element method. *IEEE Trans. Biomed. Eng.*, 41(5):455–459, 1994.

[72] Agnès Trébuchon-Da Fonseca, Christian-G Bénar, Fabrice Bartoloméi, Jean Régis, Jean-François Démonet, Patrick Chauvel, and Catherine Liégeois-Chauvel. Electrophysiological study of the basal temporal language area: A convergence zone between language perception and production networks. *Clinical Neurophysiology*, pages 1–12, Feb 2009.

[73] L. Ford and D. Fulkerson. *Flows in Networks*. Princeton University Press, 1962.

[74] M Fornasier and F Pitolli. Adaptive iterative thresholding algorithms for magnetoencephalography (meg). *Journal of Computational and Applied Mathematics*, page 10, Oct 2007.

[75] PT Fox, FM Miezin, JM Allman, DC Van Essen, and ME Raichle. Retinotopic organization of human visual cortex mapped with positron-emission tomography. *Journal of Neuroscience*, 7, 1987.

[76] W.J Freeman. Simulation of chaotic eeg patterns with a dynamic model of the olfactory system. *Biological Cybernetics*, 56:139–150, 1987.

[77] J Friedman, T Hastie, H Höfling, and R Tibshirani. Pathwise coordinate optimization. *Annals of Applied Statistics*, 1(2):302–332, Jan 2007.

[78] K Friston, L Harrison, J Daunizeau, and S Kiebel. Multiple sparse priors for the m/eeg inverse problem. *Neuroimage*, Jan 2008.

[79] K Friston, R Henson, C Phillips, and J Mattout. Bayesian estimation of evoked and induced responses. *Human brain mapping*, 27(9):722–35, Sep 2006.

[80] K.J Friston, D.E Glaser, R.N.A Henson, S Kiebel, C Phillips, and J Ashburner. Classical and bayesian inference in neuroimaging: Applications. *NeuroImage*, 16(2):484–512, 2002.

[81] K.J Friston, W Penny, C Phillips, and Kiebel. Classical and bayesian inference in neuroimaging: Theory. *NeuroImage*, 16(2):465–483, 2002.

[82] F Fylan, I Holliday, K Singh, and S Anderson. Magnetoencephalographic investigation of human cortical area v1 using color stimuli. *Neuroimage*, 6:47–57, Jan 1997.

[83] S. Gabriel, R.W. Lau, and C. Gabriel. The dielectric properties of biological tissues: Ii. measurements in the frequency range 10 hz to 20 ghz. *Physics in Medicine and Biology*, 41:2251–2269, 1996.

[84] S. Geman and D. Geman. Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6(6):721–741, 1984.

[85] C. Genovese, N. Lazar, and T. Nichols. Thresholding of statistical maps in functional neuroimaging using the false discovery rate. *NeuroImage*, 15(4):870–878, 2002.

[86] D Geselowitz. On bioelectric potentials in an homogeneous volume conductor. *Biophysics Journal*, 7:1–11, 1967.

[87] D Geselowitz. On the magnetic field generated outside an inhomogeneous volume conductor by internal volume currents. *IEEE Trans. Magn.*, 6:346–347, 1970.

[88] A P Gibson, J C Hebden, and S R Arridge. Recent advances in diffuse optical imaging. *Physics in Medicine and Biology*, 50(4):R1–43, Feb 2005.

[89] A.V. Goldberg and R.E. Tarjan. A new approach to the maximum-flow problem. *Journal of the Association for Computing Machinery*, 35(4):921–940, Oct 1988.

[90] G Golub, M Heath, and G Wahba. Generalized cross-validation as a method for choosing a good ridge parameter. *Technometrics*, Jan 1979.

[91] I Gorodnitsky, J George, and B Rao. Neuromagnetic source imaging with focuss: a recursive weighted minimum norm algorithm. *Electroencephalography and clinical Neurophysiology*, Jan 1995.

[92] I.F. Gorodnitsky and B.D. Rao. Sparse signal reconstruction from limited data using FOCUSS: A re-weighted minimum norm algorithm. *Signal Processing, IEEE Transactions on*, 45:600–616, Mar 1997.

[93] A Gramfort and M Kowalski. Improving m/eeg source localization with an inter-condition sparse prior. In *Proceedings ISBI*, Jun 2009.

[94] Alexandre Gramfort and Maureen Clerc. Low dimensional representations of MEG/EEG data using laplacian eigenmaps. In *NFSI 2007: Proceedings of the 6th International Symposium*, pages 169–172, oct 2007.

[95] G Gratton, M.R Goodman-Wood, and M Fabiani. Comparison of neuronal and hemodynamic measures of the brain response to visual stimulation : An optical imaging study. *Human Brain Mapping*, 13:13–25, 2001.

[96] D Greig, B Porteous, and A Seheult. Exact maximum a posteriori estimation for binary images. *Journal of the Royal Statistical Society, Series B*, 51(2):271–279, 1989.

[97] J Gross, J Kujala, M Hämäläinen, and L Timmermann. Dynamic imaging of coherent sources: studying neural interactions in the human brain. *Proceedings of the National Academy of Sciences*, 98(2):694–699, Jan 2001.

[98] Berger H. Über das Elektroenkephalogramm des Menschen. *Archiv für Psychiatrie und Nervenkrankheiten*, 87:527–570, 1929.

[99] Elaine T Hale, Wotao Yin, and Yin Zhang. A fixed-point continuation method for l1-regularized minimization with applications to compressed sensing. *CAAM Technical Report TR07-07*, page 45, Jul 2007.

[100] M Hämäläinen, R Hari, R Ilmoniemi, J Knuutila, and O.V Lounasmaa. Magnetoencephalography: theory, instrumentation, and applications to noninvasive studies of the working human brain. *Reviews of Modern Physics*, 65(2):413–497, 1993.

[101] M Hämäläinen and R Ilmoniemi. Interpreting magnetic fields of the brain: minimum norm estimates. *Medical and Biological Engineering and Computing*, 32(1):35–42, Jan 1994.

[102] M Hämäläinen and J Sarvas. Realistic conductivity geometry model of the human head for interpretation of neuromagnetic data. *IEEE Trans. Biomed. Eng.*, 36(2):165–171, 1989.

[103] P Hansen. Analysis of discrete ill-posed problems by means of the l-curve. *SIAM Review*, Jan 1992.

[104] R Hari and N Forss. Magnetoencephalography in the study of human somatosensory cortical processing. *Philos Trans R Soc Lond, B, Biol Sci*, 354(1387):1145–54, Jul 1999.

[105] D.A Harville. Maximum likelihood approaches to variance component estimation and to related problems. *Journal of the American Statistical Association*, 72(358):320–338, 1977.

[106] M. Hebiri. Regularization with the smooth-lasso procedure. Preprint Laboratoire de Probabilités et Modèles Aléatoires, 2008.

[107] M Hein, JY Audibert, and U von Luxburg. Graph Laplacians and their Convergence on Random Neighborhood Graphs. *The Journal of Machine Learning Research*, 8:1325–1370, 2007.

[108] Ming-Xiong Huang, Anders M Dale, Tao Song, Eric Halgren, Deborah L Harrington, Igor Podgorny, Jose M Canive, Stephen Lewis, and Roland R Lee. Vector-based spatial-temporal minimum l1-norm solution for meg. *Neuroimage*, 31(3):1025–37, Jul 2006.

[109] G Huiskamp, M Vroeijenstijn, R Dijk, G Wieneke, and A Huffelen. The need for correct realistic geometry in the inverse EEG problem. *IEEE Trans. on Biomed. Engin.*, 46(11):1281–1287, 1999.

[110] Chang-Hwan Im, Arvind Gururajan, Nanyin Zhang, Wei Chen, and Bin He. Spatial resolution of eeg cortical source imaging revealed by localization of retinotopic organization in human primary visual cortex. *J Neurosci Methods*, 161(1):142–54, Mar 2007.

[111] Bancaud J, Talairach J, Bonis A, Schaub C, Szikla G, and Morel P et al. *La stéréoélectroencephalographie dans l'épilepsie: informations neurophysiopathologiques apportées par l'investigation fonctionelle stéréotaxique.* Paris, Masson, 1965.

[112] L. Jacob, G. Obozinski, and J.-P. Vert. Group Lasso with Overlap and Graph Lasso. In *ICML'09 Proceedings of the 26th international conference on Machine learning*, 2009.

[113] Ben Jansen and Vincent Rit. Electroencephalogram and visual evoked potential generation in a mathematical model of coupled cortical columns. *Biol. Cybern.*, 73:357–366, 1995.

[114] P Jaskowski and R Verleger. Amplitudes and latencies of single-trial ERP's estimated by a maximum-likelihood method. *IEEE Transactions on Biomedical Engineering*, 46(8):987–993, Aug 1999.

[115] H. H. Jasper. The ten-twenty electrode system of the International Federation. *Electroencephalography and Clinical Neurophysiology*, 10:371–375, 1958.

[116] R Jenatton, J-Y Audibert, and F Bach. Structured variable selection with sparsity-inducing norms. Technical report, WILLOW (INRIA Rocquencourt), Imagine, 2009.

[117] K Jerbi, Sylvain Baillet, J.C Mosher, G Nolte, L Garnero, and R.M Leahy. Localization of realistic cortical activity in meg using currentmultipoles. *Neuroimage*, 22(2):779–793, 2004.

[118] K Jerbi, C Mosher, S Baillet, and R.M Leahy. On meg forward modelling using multipolar expansions. *Physics in Medicine and Biology*, 47:523–555, 2002.

[119] K. Jerbi, J.C. Mosher, S. Baillet, and R.M. Leahy. On MEG forward modelling using multipolar expansions. *Phys. Med. Biol.*, 47:523–555, 2002.

[120] Moreau J.J. Proximité et dualité dans un espace hilbertien. *Bull. Soc. Math. France.*, 93:273–299, 1965.

[121] E.G Jones and A Peters. *Cerebral cortex, functional properties of cortical cells*, volume 2. Plenum Press, 1984.

[122] O Juan and Y Boykov. Active graph cuts. In *Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on*, volume 1, pages 1023–1029, 2006.

[123] T.P. Jung, S. Makeig, M. Westerfield, J. Townsend, E. Courchesne, and T.J. Sejnowski. Analysis and visualization of single-trial event-related potentials. *Human Brain Mapping*, 14:166–185, 2001.

[124] E.R Kandel, J.H Schwartz, and T.M Jessel. *Principles of Neural Science*. McGraw-Hill Education, 2000.

[125] C Koch. *Biophysics of Computation: Information Processing in Single Neurons*. Oxford University Press, USA, 1999.

[126] V. Kolmogorov and R. Zabih. Computing visual correspondence with occlusions using graph cuts. In *Computer Vision, 2001. ICCV 2001. Proceedings. Eighth IEEE International Conference on*, volume 2, pages 508–515, 2001.

[127] V Kolmogorov and R Zabih. What energy functions can be minimized via graph cuts? *IEEE Transactions on Pattern Analysis and Machine*, Jan 2004.

[128] M Kowalski and A Gramfort. A priori par normes mixtes pour les problèmes inverses: Application à la localisation de sources en M/EEG. In *Proceedings GRETSI*, Jun 2009.

[129] M Kowalski and B Torrésani. Sparsity and persistence: mixed norms provide simple signals models with dependent coefficients. *Signal, Image and Video Processing*, 2008.

[130] M Kutas, G McCarthy, and E Donchin. Augmenting mental chronometry: the P300 as a measure of stimulus evaluation time. *Science*, 197:792–795, Aug 1977.

[131] J Kybic, M Clerc, T Abboud, O Faugeras, R Keriven, and T Papadopoulo. A common formalism for the integral formulations of the forward eeg problem. *IEEE Transactions on Medical Imaging*, 24(1):12–28, 2005.

[132] J Kybic, M Clerc, O Faugeras, R Keriven, and T Papadopoulo. Generalized head models for meg/eeg: boundary element method beyond nested volumes. *Phys. Med. Biol.*, 51:1333–1346, 2006.

[133] J.-P. Lachaux, E. Rodriguez, Jacques Martinerie, and Francisco Varela. Measuring phase-synchrony in brain signals. *Human Brain Mapping*, 8(4):194– –208, Nov 1999.

[134] D Lange, H Pratt, and G Inbar. Modeling and estimation of single evoked brain potential components. *IEEE Transactions on Biomedical Engineering*, 44(9):791–799, Sep 1997.

[135] J Lefevre and S Baillet. Optical flow and advection on 2-riemannian manifolds: A common framework. *IEEE transactions on Pattern Analysis and Machine Intelligence*, 30(6):1081–1092, 2008.

[136] J Lefevre, G Obozinski, and S Baillet. Imaging brain activation streams from optical flow computation on 2-riemannian manifolds. *IPMI 2007, Lecture notes in Computer Science*, 4587:470–481, Jan 2007.

[137] R. Lehoucq, D. Sorensen, and D. Yang. ARPACK users' guide: Solution of large-scale eigenvalue problems with implicitly restarted Arnoldi methods. *SIAM Publications, Philadelphia*, Jan 1998.

[138] M. Leventon, E. Grimson, and O. Faugeras. Statistical Shape Influence in Geodesic Active Contours. In *CVPR*, page 316–323, 2000.

[139] Y Li. A globally convergent method for l_p problems. *SIAM Journal on Optimization*, 3(3):609–629, 1993.

[140] Fa-Hsuan Lin, Thomas Witzel, Seppo P. Ahlfors, Steven M. Stufflebeam, John W. Belliveau, and Matti S. Hämäläinen. Assessing and improving the spatial accuracy in meg source localization by depth-weighted minimum-norm estimates. *NeuroImage*, 31(1):160–171, May 2006.

[141] Kowalski M. Sparse regression using mixed norms. *Applied and Computational Harmonic Analysis*, page In press, 2009.

[142] David J. C. Mackay. Bayesian interpolation. *Neural Computation*, 4:415–447, 1992.

[143] S. Mallat. *A Wavelet Tour of Signal Processing*. Academic Press, 1998.

[144] S Mallat and Z Zhang. Matching pursuit with time-frequency dictionaries. *IEEE Trans. on Signal Processing*, 41(12):3397–3414, 1993.

[145] K. Matsuura and Y. Okabe. Selective minimum-norm solution of the biomagnetic inverse problem. *IEEE Trans Biomed Eng*, 42(6):608–615, June 1995.

[146] J Mattout, C Phillips, W Penny, and M Rugg. Meg source localization under multiple constraints: An extended bayesian framework. *Neuroimage*, Jan 2006.

[147] C McGillem, J Aunon, and C Pomalaza. Improved waveform estimation procedures for event-related potentials. *IEEE Transactions on Biomedical Engineering*, 32(6):371–379, Jun 1985.

[148] J Meijs, O Weier, M Peters, and A van Oosterom. On the numerical accuracy of the boundary element method. *IEEE Trans. Biomed. Eng.*, 36:1038–1049, 1989.

[149] J. W. H. Meijs and M. Peters. The EEG and MEG using a model of eccentric spheres to describe the head. *IEEE Transactions on Biomedical Engineering*, 34:913–920, 1987.

[150] S Meunier, L Garnero, A Ducorps, and Maziï¿½es. Human brain mapping in dystonia reveals both endophenotypic traits and adaptative reorganization. *Annals of Neurology*, 50:521–527, 2001.

[151] F Moradi, L.C Liu, K Cheng, R.A Waggoner, K Tanaka, and A.A Ioannides. Consistent and precise localization of brain activity in human primary visual cortex by meg and fmri. *NeuroImage*, 18:595–609, 2003.

[152] J Mosher, S Baillet, and R Leahy. Equivalence of linear approaches in bioelectromagnetic inverse solutions. *Statistical Signal Processing*, Jan 2003.

[153] J Mosher and R Leahy. Source localization using recursively applied and projected (rap) music. *Signal Processing*, 47(2):332–339, Jan 1999.

[154] J Mosher, P Lewis, and R Leahy. Multiple dipole modeling and localization from spatio-temporal megdata. *Biomedical Engineering*, Jan 1992.

[155] John Mosher, Richard Leahy, and Paul Lewis. Eeg and meg: Forward solutions for inverse methods. *IEEE Transactions on Biomedical Engineering*, 46(3):245–259, 1999.

[156] John Mosher, Paul Lewis, and Richard Leahy. Multiple dipole modeling and localization from spatio-temporal meg data. *IEEE Transactions on Biomedical Engineering*, 39(6):541–553, 1992.

[157] V.B Mountcastle. Modality and topographic properties of single neurons of cat's somatosensory cortex. *Journal of Neurophysiology*, 20:408–434, 1957.

[158] D Mumford and J Shah. Optimal approximations by piecewise smooth functions and associated variational problems. *Comm. Pure Appl. Math*, Jan 1989.

[159] S Murakami and Y Okada. Contributions of principal neocortical neurons to magnetoencephalography and electroencephalography signals. *The Journal of Physiology*, 575(3):925–936, 2006.

[160] Radford M. Neal. *Bayesian Learning for Neural Networks (Lecture Notes in Statistics)*. Springer, 1 edition, August 1996.

[161] Y Nesterov. Gradient methods for minimizing composite objective function. CORE Discussion Papers 2007076, Université catholique de Louvain, Center for Operations Research and Econometrics (CORE), Sep 2007.

[162] T. Noesselt, S. A. Hillyard, M. G. Woldorff, A. Schoenfeld, T. Hagner, L. Jancke, C. Tempelmann, H. Hinrichs, and H. J. Heinze. Delayed striate cortical activation during spatial attention. *Neuron*, 35:575–587, 2002.

[163] John Nolte. *The human brain: an introduction to its functional anatomy*. Mosby-Year Book, 3 edition, 1993.

[164] Jean-Claude Nédélec. *Acoustic and Electromagnetic Equations*. Springer Verlag, 2001.

[165] T. F. Oostendorp and A. van Oosterom. Source parameter estimation in inhomogeneous volume conductors of arbitrary shape. *IEEE Trans. Biomed. Eng.*, BME-36:382–391, 1989.

[166] W Ou, M Hämaläinen, and P Golland. A distributed spatio-temporal eeg/meg inverse solver. *Neuroimage*, 44:932–946, 2009.

[167] S. E. Palmer. *Vision Science-Photons to Phenomenology*. MIT Press, Cambridge, MA, 1999.

[168] D Pantazis, T Nichols, S Baillet, and R Leahy. A comparison of random field theory and permutation methods for the statistical analysis of meg data. *Neuroimage*, Jan 2005.

[169] D. Pantazis, Thomas E Nichols, Sylvain Baillet, and R.M. Leahy. Spatiotemporal localization of significant activation in meg using permutation tests. *Inf Process Med Imaging*, 18:512–523, Jul 2003.

[170] Théodore Papadopoulo and Sylvain Vallaghé. Implicit meshing for finite element methods using levelsets. In *Proceedings of MMBIA 07*, 2007.

[171] R Pascual-Marqui. Standardized low resolution brain electromagnetic tomography (sloreta): technical details. *Methods Find. Exp. Clin. Pharmacology*, 24(D):5–12, Jan 2002.

[172] R. D. Pascual-Marqui, C. M. Michel, and D. Lehman. Low resolution electromagnetic tomography: A new method for localizing electrical activity of the brain. *Psychophysiology*, 18:49–65, 1994.

[173] M Pastor, J Artieda, J Arbizu, and M Valencia. Human cerebral activation during steady-state visual-evoked responses. *Journal of neuroscience*, Jan 2003.

[174] W. Penfield and T. Rasmussen. *The Cerebral Cortex of Man: A Clinical Study of Localization of Function*. Macmillan, 1950.

[175] Alan Peters and Edward G. Jones, editors. *Cellular Components of the Cerebral Cortex*, volume 1 of *Cerebral Cortex*. Plenum, New York, 1984.

[176] C Phillips. *Source estimation in EEG*. PhD thesis, University de Liege, Belgium, 2000.

[177] C Phillips, J Mattout, M Rugg, and P Maquet. An empirical bayesian solution to the source reconstruction problem in eeg. *Neuroimage*, Jan 2005.

[178] C Phillips, M Rugg, and K Friston. Anatomically informed basis functions for eeg source localization: Combining functional and .... *Neuroimage*, Jan 2002.

[179] J Picard and H Ratliff. Minimum cuts and related problems. *Networks*, 5(4):357–370, Jan 1975.

[180] B Presnell, B Turlach, and M Osborne. A new approach to variable selection in least squares problems. *IMA Journal of Numerical Analysis*, 20:389–404, 2000.

[181] R. Q. Quiroga and H. Garcia. Single-trial event-related potentials with wavelet denoising. *Clinical Neurophysiology*, 114(2):376–290, 2003.

[182] Ramírez R and Makeig S. Neuroelectromagnetic source imaging using multiscale geodesic neural bases and sparse bayesian learning. In *Human Brain Mapping*, 2006.

[183] G Rager and W Singer. The response of cat visual cortex to flicker stimuli of variable frequency. *The European journal of neuroscience*, 10(5):1856–1877, 1998.

[184] M. Raichle. A brief history of human brain mapping. *Trends in Neurosciences*, December 2008.

[185] P.O. Ranta-aho, A.S. Koistinen, J.O. Ollikainen, J.P. Kaipio, J. Partanen, and P.A. Karjalainen. Single-trial estimation of multichannel evoked-potential measurements. *IEEE Transactions on Biomedical Engineering*, 50(2):189–196, 2003.

[186] B Rao, K Engan, S Cotter, J Palmer, and K Kreutz-Delgado. Subset selection in noise based on diversity measure minimization. *Signal Processing*, Jan 2003.

[187] D. Regan. *Human Brain Electrophysiology: Evoked Potentials and Evoked Magnetic Fields in Science and Medicine*. Elsevier, 1989.

[188] R.T Rockafellar. *Convex analysis*. Princeton University Press, 1970.

[189] L Rudin, S Osher, and E Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D*, 60:259–268, 1992.

[190] Francesco Di Russo, Antígona Martínez, Martin I. Sereno, Sabrina Pitzalis, and Steven A. Hillyard. Cortical sources of the early components of the visual evoked potential. *Humain Brain Mapping*, 15:95–111, 2002.

[191] Francesco Di Russo, Sabrina Pitzalis, Teresa Aprile, Grazia Spitoni, Fabiana Patria, Alessandra Stella, Donatella Spinelli, and Steven A Hillyard. Spatiotemporal analysis of the cortical sources of the steady-state visual evoked potential. *Human brain mapping*, 28(4):323–334, Apr 2007.

[192] Jukka Sarvas. Basic mathematical and electromagnetic concepts of the biomagnetic inverse problem. *Phys. Med. Biol.*, 32(1):11–22, 1987.

[193] M Scherg and D Von Cramon. Two bilateral sources of the late AEP as identified by a spatio-temporal dipole model. *Electroencephalography and Clinical Neurophysiology*, 62:32–44, 1985.

[194] K. Sekihara, S. Nagarajan, D. Poeppel, and Y. Miyashita. Reconstructing spatio-temporal activities of neural sources from magnetoencephalographic data using a vector beamformer. In *ICASSP '01: Proceedings of the Acoustics, Speech, and Signal Processing, 2001. on IEEE International Conference*, pages 2021–2024, Washington, DC, USA, 2001. IEEE Computer Society.

[195] M.I Sereno, A.M Dale, J.B Reppas, and Kwong. Borders of multiple visual areas in human revealed by functional magnetic resonance imaging. *Science*, pages 889–893, 1995.

[196] R Shapley and J Victor. The effect of contrast on the transfer properties of cat retinal ganglion cells. *The Journal of Physiology*, 285(1):275–298, 1978.

[197] D Sharon, M Hämäläinen, R Tootell, and E Halgren. The advantage of combining meg and eeg: Comparison to fmri in focally stimulated visual cortex. *Neuroimage*, 36:1225–1235, Mar 2007.

[198] X Shen and F Meyer. Low-dimensional embedding of fMRI datasets. *Neuroimage*, 41(3):886–902, Jan 2008.

[199] P. Suffczynski, S. Kalitzin, G. Pfurtscheller, and FH Lopes da Silva. Computational model of thalamo-cortical networks: dynamical control of alpha rhythms in relation to focal attention. *International Journal of Psychophysiology*, 43(1):25–40, 2001.

[200] J Talairach, J Bancaud, and G Szikla. Approche nouvelle de la neurochirugie de l'epilepsie. méthodologie stérérotaxique et résultats thérapeutiques. *Neurochirurgie*, 20:1–240, 1974.

[201] A Tarantola. Popper, bayes and the inverse problem. *Nature Physics*, 2, Aug 2006.

[202] R Tibshirani. Regression shrinkage and selection via the Lasso. *J.R. Statist. Soc.*, 58(1):267–288, 1996.

[203] A.N Tikhonov and V.Y Arsenin. *Solutions of Ill-Posed Problems*. Winston & Sons, Washington, 1977.

[204] R Tootell, N Hadjikhani, J Mendola, S Marrett, and A Dale. From retinotopy to recognition: fmri in human visual cortex. *Trends in Cognitive Sciences*, 2(5):174–183, 1998.

[205] R. Tootell, E. Switkes, M. Silverman, and S. Hamilton. Functional anatomy of the macaque striate cortex. ii. retinotopic organization. *Journal of neuroscience*, 8(5):1531–1568, 1988.

[206] W Truccolo, K H Knuth, A Shah, S L Bressler, C E Schroeder, and M Ding. Estimation of single-trial multicomponent ERPs: Differentially variable component analysis (dVCA). *Biological Cybernetics*, 89(6):426–438, Dec 2003.

[207] P D Tuan, J Möcks, W Köhler, and T Gasser. Variable latencies of noisy signals: Estimation and testing in brain potential data. *Biometrika*, 74(3):525–533, 1987.

[208] M Uusitalo and R Ilmoniemi. Signal-space projection method for separating MEG or EEG into components. *Medical and Biological Engineering and Computing*, 35:135–140, Jan 1997.

[209] K Uutela, M Hamalainen, and R Salmelin. Global optimization in the localization of neuromagnetic sources. *IEEE Transactions on Biomedical Engineering*, 45(6):716–723, June 1998.

[210] Pedro A Valdés-Sosa, Mayrim Vega-Hernández, José Miguel Sánchez-Bornot, Eduardo Martínez-Montes, and María Antonieta Bobes. Eeg source imaging with spatio-temporal tomographic nonnegative independent component analysis. *Human Brain mapping*, 30(6):1898–910, Jun 2009.

[211] Sylvain Vallaghé. *EEG and MEG forward modeling : computation and calibration*. PhD thesis, Universite de Nice-Sophia Antipolis, 2008.

[212] E van den Berg and M Friedlander. Probing the pareto frontier for basis pursuit solutions. *Department of Computer Science*, Jan 2008.

[213] S. Vanni, J. Warnking, M. Dojat, C. Delon-Martin, J. Bullier, and C. Segebarth. Sequence of pattern onset responses in the human visual areas: an fMRI constrained VEP source analysis. *NeuroImage*, 21(3):801–817, 2004.

[214] S Vanni, J Warnking, M Dojat, C Delon-Martin, J Bullier, and C Segebarth. Sequence of pattern onset responses in the human visual areas: an fmri constrained vep source analysis. *NeuroImage*, 21:801–817, 2004.

[215] B Van Veen, W Van Drongelen, M Yuchtman, and A Suzuki. Localization of brain electrical activity via linearly constrained minimum variance spatial filtering. *Biomedical Engineering*, 44(9):867—880, Jan 1997.

[216] J. Vernon Odom, M. Bach, C. Barber, M. Brigell, M.F. Marmor, and A.P. Tormene. Visual evoked potentials standard. *Documenta Ophthalmologica*, 108:115–123, 2004.

[217] J Vrba and E Robinson. Signal processing in magnetoencephalography. *Methods*, 25(2):249–271, Oct 2001.

[218] G Wahba. Practical approximate solutions to linear operator equations when the data are noisy. *SIAM Journal on Numerical Analysis*, Jan 1977.

[219] B Wandell, S Dumoulin, and A Brewer. Visual field maps in human cortex. *Neuron*, 56(2):366–383, Oct 2007.

[220] J.-Z. Wang, S.J. Williamson, and L. Kaufman. Magnetic source images determined by a lead-field analysis: the unique minimum-norm least-squares estimation. *Biomedical Engineering, IEEE Transactions on*, 39(7):665–675, July 1992.

[221] Z. Wang, A. Maier, D.A. Leopold, N.K. Logothetis, and H. Liang. Single-trial evoked potential estimation using wavelets. *Computers in Biology and Medicine*, 37(4):463–473, Apr 2007.

[222] J Warnking. *Délinéation des aires visuelles rétinotopiques chez l'homme par IRM fonctionnelle*. PhD thesis, Université Joseph Fourier-Grenoble I, 2002.

[223] J Warnking, M Dojat, A Guerin-Dugue, C Delon-Martin, S Olympieff, N Richard, A Chehikian, and C Segebarth. fmri retinotopic mapping - step by step. *NeuroImage*, 17:1665–1683, 2002.

[224] Pierre Weiss. *Algorithmes rapides d'optimisation convexe. Applications à la reconstruction d'images et à la détection de changements.* PhD thesis, Université de Nice Sophia-Antipolis, Novembre 2008.

[225] P. Welch. The use of fast Fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms. *Audio and Electroacoustics, IEEE Transactions on*, 15(2):70–73, Jun 1967.

[226] F Wendling, J.J Bellanger, F Bartolomei, and P Chauvel. Relevance of nonlinear lumped-parameter models in the analysis of depth-eeg epileptic signals. *Biological Cybernetics*, 83:367–378, 2000.

[227] D Wipf and S Nagarajan. A unified bayesian framework for meg/eeg source imaging. *Neuroimage*, 44(3):947–966, Feb 2009.

[228] Adrien Wohrer and Pierre Kornprobst. *Virtual Retina* : A biological retina model and simulator, with contrast gain control. *Journal of Computational Neuroscience*, 26(2):219–249, 2009.

[229] C. H. Wolters, A. Anwander, X. Tricoche, D. Weinstein, M. A. Koch, and R. MacLeod. Influence of tissue conductivity anisotropy on EEG/MEG field and return current computation in a realistic head model: A simulation and visualization study using high-resolution finite element modeling. *NeuroImage*, 3:813–826, 2006.

[230] C.D. Woody. Characterization of an adaptive filter for the analysis of variable latency neuroelectrical signals. *Medical and Biological Engineering*, 5:539–553, 1967.

[231] Z. Wu and R. Leahy. An optimal graph theoretic approach to data clustering: Theory and its application to image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15(11):1101–1113, 1993.

[232] Ning Xu, Ravi Bansal, and Narendra Ahuja. Object segmentation using graph cuts based active contours. *Computer Vision and Pattern Recognition, IEEE Computer Society Conference on*, 2:46, 2003.

[233] Alper Yilmaz, Omar Javed, and Mubarak Shah. Object tracking: A survey. *ACM Comput. Surv.*, 38(4), 2006.

[234] D. Yoshor, W. H. Bosking, G. M. Ghose, and J. H. Maunsell. Receptive fields in human visual cortex mapped with surface electrodes. *Cereb Cortex*, 17(10):2293–2302, October 2007.

[235] M Yuan and Y Lin. Model selection and estimation in regression with grouped variables. *Journal of the Royal Statistical Society*, Jan 2006.

[236] Benjamin W. Zeff, Brian R. White, Hamid Dehghani, Bradley L. Schlaggar, and Joseph P. Culver. Retinotopic mapping of adult human visual cortex with high-density diffuse optical tomography. *Proceedings of the National Academy of Sciences*, 104(29):12169–12174, 2007.

[237] LH Zetterberg, L. Kristiansson, and K. Mossberg. Performance of a model for a local neuron population. *Biological Cybernetics*, 31(1):15–26, 1978.

[238] Zhi Zhang. A fast method to compute surface potentials generated by dipoles within multilayer anisotropic spheres. *Phys. Med. Biol.*, 40:335–349, 1995.

[239] H Zou and T Hastie. Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society Series B*, Jan 2005.