

CHALMERS



Development of an Analysis Pipeline for Magnetoencephalography Measurement Data Using the MNE Software Package

Master's Thesis in Biomedical Engineering

MOA PETER

Department of Signals & Systems
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden 2015

REPORT NO. EX013/2015

**Development of an Analysis Pipeline
for Magnetoencephalography
Measurement Data using the MNE
Software Package**

Moa Peter
February 2015

Development of an Analysis Pipeline for Magnetoencephalography Measurement
Data using the MNE Software Package

Moa Peter

©Moa Peter, 2015
Technical Report no. EX013/2015
Department of Signals and Systems
Chalmers University of Technology
412 96 Göteborg
Sweden +46 (0)31-772 1000

Abstract

Magnetoencephalography (MEG) is a non-invasive, functional neuroimaging method used to map the neural activity within the brain by measuring the magnetic fields generated by neuronal currents. One important benefit with MEG is its sub-millisecond temporal resolution, which is faster than the most rapid neuro-dynamics observed to date. This enables better understanding of the mechanisms of the brain as compared to common functional neuroimaging methods such as functional magnetic resonance imaging (fMRI with a temporal resolution of about 1 s). There are many challenges in MEG data analysis, e.g., the data is often affected by environmental noise and physiological artifacts of which the most salient are caused by eye blinks, heartbeats, and muscle activity. It is important to put effort and thought into the MEG data analysis pipeline to be able to present the measurement data in a reliable and neurophysiologically correct way for medical experts to interpret. Furthermore, a processing pipeline needs to be tailored for each MEG study based on variables such as the hypothesis under test, the measurement equipment used, and the experimental protocol. The purpose of this thesis work was to develop a pipeline for MEG data analysis with focus on preprocessing of measurement data and analysis on sensor level (studying the magnetic fields rather than the computed neuronal activity). The pipeline was developed using the MNE software package and optimized for analyzing event related fields from somatosensory stimuli. An explorative and iterative method was used that was based on analyzing four MEG data sets from pilot measurements within an ongoing medical research project. Furthermore, feedback from medical and technical experts enabled evaluation and selection of processing methods within the pipeline. Analysis methods discussed in this thesis are, e.g., visual inspection of measurement data, filtering and averaging for noise reduction, as well as artifact removal methods including signal space separation, signal space projection, and independent component analysis.

Acknowledgements

There are so many people who have helped and supported me during this thesis work, for which I am extremely grateful. I would like to take this opportunity to mention a few of them.

First of all I would like to thank my supervisor and examiner Justin Schneiderman for giving me the opportunity to do this thesis work, for his support and guidance, and for being an inspiration. I would also like to extend a special thank you to Bushra Riaz, with whom I've worked closely throughout this process. Without Justin and Bushra, this thesis wouldn't have been possible.

Mikael Elam has allowed me to use measurement data from his research for the development of the pipeline and has contributed enormously with his medical expertise, improving my understanding of the medical aspects of MEG data analysis.

Daniel Lundqvist and Stephen Whitmarsh have been extremely kind and helpful. They have provided insight into how the MEG system at NatMEG works and have helped with the MEG data analysis, giving invaluable feedback on the analysis pipeline during the development process.

I would also like to thank all of my colleagues at MedTech West for providing an inspiring and pleasant work environment.

Last but not least I want to express my gratitude to my friends and family for always supporting me and being there when I'm in need of feedback, pep talks, hugs and chocolate.

Moa

Contents

1	Introduction	1
1.1	Background	1
1.1.1	Magnetoencephalography	2
1.1.2	MedTech West	2
1.1.3	Arousal Project	3
1.2	Motivation	3
1.3	Aim	4
1.4	Scope	4
1.5	Thesis outline	4
2	MEG Theory	6
2.1	Neuroanatomy and Biomagnetism	7
2.1.1	Macroscopic brain	7
2.1.2	Microscopic brain	8
2.1.3	Neuromagnetism	9
2.2	MEG Sensors	10
3	Materials and Method	12
3.1	Work Procedure	12
3.2	Data Acquisition	13
3.2.1	Equipment	13
3.2.2	Subjects	13
3.2.3	Experimental protocol	13
3.3	MNE Software Package	14
3.4	Developing the Analysis Pipeline	15
3.4.1	Preprocessing	16
3.4.2	Averaging	17
3.4.3	Artifact and Noise Reduction	17
3.4.4	Statistical analysis	20
3.4.5	Forward Solution	20

3.4.6	Inverse Solution	21
3.4.7	Script Editing	21
3.4.8	Testing on New Data Sets	22
4	Results and Discussion	23
4.1	Analysis results	23
4.1.1	Preprocessing	23
4.1.2	Averaging and Artifact Removal	26
4.1.3	Statistical Analysis	31
4.1.4	Source Level	33
4.2	Testing on New Data Sets	34
4.3	Analysis Pipeline	35
5	Conclusion	38
5.1	Limitations	38
5.2	Future Aspects	39
	Bibliography	42
A	Python Script	43
B	Generated Report	51

1

Introduction

1.1 Background

The workings of the human body and mind has fascinated human kind for centuries. The earliest medical records of the brain, describing parts of its structures and relating brain injuries to functions in other parts of the body, are found in the *Edwin Smith Surgical Papyrus*, dating back to Ancient Egypt, the 17th Century B.C.[1]. The Ancient Greek Alcmaeon of Croton is, however, thought to be the first one stating that the brain is the organ in which the human mind is located[2]. Today, the brain is investigated in many different fields of research, stretching from philosophy to psychology and neurology and during the last decades even technology, with research within artificial neural networks and the Human Brain Project, a largely EU funded 10-year project aiming to simulate a human brain on supercomputers[3]. Despite the large amount of research done on human brain anatomy and physiology, there are still vast knowledge gaps when it comes to understanding the workings of the human brain. A deeper understanding of the structure and the functions of the human brain can be directly beneficial medically in terms of investigating brain disorders, and basic research on healthy human brains can help giving insight into our understanding of ourselves as well as providing information that can be used in future clinical applications.

Since it is impossible to observe the anatomy and function of the brain by visual inspection of a human being without opening his/her skull, techniques for non-invasive neuroimaging have been developed. There are today a vast number of different methods, which all have their own benefits and drawbacks in terms of costs, results and ease of use. There are purely structural (also called anatomical) neuroimaging methods like computed tomography (CT), based on X-rays, and magnetic resonance imaging (MRI), utilising a combination of strong magnetic fields and radiowaves to form structural images of the brain. Functional neuroimaging methods describing the activity within the brain also exist. Some examples are functional MRI (fMRI), based on blood-oxygenation-levels, positron emission tomography (PET), based on measurements of emissions of radioactively labelled chemicals injected into the bloodstream, and electro- and magne-

toencephalography (EEG and MEG), measuring potentials and magnetic fields caused by neuronal currents, respectively. When evaluating the benefits and drawbacks of functional methods, it is important to consider both spatial and temporal resolution of the images. The spatial resolutions of PET, fMRI and MEG are all on the level of a few mm, with fMRI having the capacity of creating images with 1 mm resolution. The temporal resolution of PET is tens of seconds[4], for fMRI it's around one second[4] and MEG has a temporal resolution on a sub-millisecond level[4]. A high temporal resolution enables a more accurate picture of the brain processing (since the neuronal activity can be on the ms scale), which in turn can facilitate the understanding of the mechanisms of the normal brain as well as different brain disorders.

1.1.1 Magnetoencephalography

Magnetoencephalography (MEG) is a functional neuroimaging method utilising measurements of the magnetic fields outside the head that are generated by activity within the brain in form of neuronal currents. MEG is completely non-invasive and used without the application of external fields (as compared to MRI/fMRI), ionising radiation (as compared to CT) or the injection of radioactive tracers in the bloodstream (as compared to PET) - MEG "passively" measures effects of the brain activity. Since MEG is measuring the fields from the neuronal currents directly instead of being dependent of hemodynamic changes related to the neuronal activity (like fMRI and PET), MEG has the great benefit of a very high temporal resolution on the millisecond[5] or even sub-millisecond[4] level. The spatial resolution of MEG is varying depending on factors as data handling algorithms and noise levels, but is approximated as 5 mm[4] going down to 2-3 mm under favourable circumstances[6]. The sources of the MEG signals, the neuronal currents giving rise to the magnetic fields measured by MEG, are the same sources giving rise to the electric potentials measured on the scalp by the more common technique electroencephalography (EEG). Both methods have their limitations and benefits as compared to each other, with slightly different sensitivities depending on the direction of the neuronal current. Some advantages that MEG has over EEG are that the magnetic fields are less distorted by skin and skull impedance and head inhomogeneities than electric potentials and MEG also has a short setup time without the need of attaching an array of electrodes to the subject's head, making it convenient to use for subject as well as researcher.

MEG is used in clinical applications such as identifying epileptic foci[7], presurgical mapping e.g. identifying the primary somatosensory cortex and primary motor cortex[4] and shows promises in stroke recovery monitoring[7]. MEG is also used in research related to, e.g., cognition and perception, schizophrenia, autism, stuttering, traumatic brain injuries etc.[4]

1.1.2 MedTech West

This thesis work is done at MedTech West, a joint venture founded by Chalmers University of Technology, the University of Gothenburg (UGOT), University of Borås, Västra

Götalandsregionen and Sahlgrenska University Hospital. MedTech West aims to increase the quality and quantity of research in the field of medical technology by being a collaborative platform for research and development as well as education and evaluation; linking academia to the healthcare sector and industry. One of MedTech West's main research areas is sensor technology, including a focal MEG system based on high T_C -SQUIDS (high critical temperature superconducting quantum interference devices) currently under development. The aim with the focal MEG system is to improve sensitivity and spatial resolution of the measurements and at the same time maintaining lower operational costs as compared to state-of-the-art in MEG systems[8]. Apart from the sensor technology development, MedTech West is also involved in medical research projects in which MEG measurements and the analysis of MEG measurement data is vital. One of the medical research projects, which this thesis is coupled to, is done in collaboration with (amongst others) professor Mikael Elam at the Department of Clinical Neurosciences, UGOT and currently goes under the name *Arousal Project*. The aim of the project is to use a neuroscience approach to study cardiovascular diseases based on previous studies revealing two different sympathetic nerve response profiles as a reaction to environmental stress[9].

The planned MEG measurements (as well as the already conducted pilot measurements) within the arousal project will be performed at NatMEG, the Swedish National Facility for Magnetoencephalography, at Karolinska Institutet, Stockholm. The same goes for the other ongoing medical projects MedTech West is involved in.

1.1.3 Arousal Project

Previous research has shown two different sympathetic nerve response profiles as a reaction to arousal[9][10]. In the most recent article written by Donadio et al. 2012[9], approximately 50% of subjects, called responders, showed a reduction in muscle sympathetic nerve activity (MSNA) as a reaction to an arousal stimulus in the form of an electric current pulse delivered to the index finger of the subject. This reduction was not evident in the remaining 50% of subjects, called non-responders. These arousal response profiles are linked to mental stress MSNA and blood pressure responses, and it is concluded that MSNA responses to arousal predicts the MSNA and blood pressure response to mental stress[9].

The central hypothesis of the ongoing research is that the two profiles are associated with different risks of hypertension and thereby diseases related to this condition. An ambition within the project is to study the brain processing related to these response profiles using MEG. This is to get a better understanding of the neurophysiological differences between profiles and replace microneurography, the invasive and rather difficult method previously used to study the sympathetic nerve activity.

1.2 Motivation

When using MEG as a neuroimaging method, processing and analyzing the measurement data is a vital step. Processing the raw measurement data is needed to be able to model

the brain activity at all, and it is extremely important to put effort and thought into the MEG data analysis pipeline to be able to present the measurement data in a reliable, neurophysiologically correct way for the neurologists to interpret. Although there are standard processing steps performed in MEG data analysis, a processing pipeline needs to be tailored for each MEG study, based on variables like the hypothesis under test, the measurement equipment used and the experimental protocol[5].

1.3 Aim

This thesis work is focused on developing a MEG data analysis pipeline using the MNE software package to facilitate and improve the interpretation of MEG measurement data. The pipeline is developed for, and therefore first and foremost applicable to, the ongoing Arousal Project. However, the idea is to broaden the knowledge regarding MEG analysis at MedTech West, and utilizing the knowledge in the other medical research projects as well as for the analysis of data coming from the focal MEG system currently under development. The aim is to develop a pipeline including analysis scripts as well as a concise “user manual” describing the different analysis steps to take and scripts to be used.

1.4 Scope

The analysis pipeline is developed using MEG data from the NatMEG facility’s Elekta Neuromag TRIUX system. Different softwares for MEG data analysis exist, but the analysis pipeline development in this thesis was confined to the MNE software package, one of the two softwares for MEG measurement data processing used and taught at the NatMEG facility. The reason for choosing MNE over the other software is the fact that the MNE software was developed in parallel with the development of the Elekta MEG systems, and is therefore specifically adapted to it.

The analysis protocol was developed for analyzing evoked fields from somatosensory stimuli, based on four data sets from the Arousal Project. The focus of this thesis has been on single-subject sensor level analysis, i.e. analysing the measured magnetic fields rather than the neuronal sources of these fields, although source level analysis is briefly discussed. The ambition has been to understand how to use the MNE software for MEG data analysis in a simple and user friendly way and create a simple and user friendly pipeline. Therefore, the methods and processing steps available in the MNE software were studied, tested and compared without exhaustively investigating the mathematical basis and source codes of them.

1.5 Thesis outline

This chapter (*Introduction*) provides a background to the area of neuroimaging in general and MEG in particular as well as previous research connected to this thesis work.

The motivation, aim and scope of this thesis work is also presented herein.

Chapter 2, *MEG Theory*, lays the theoretical basis of MEG, touching on both the biological sources of the measured fields and the measurement system itself.

Chapter 3, *Materials and Method*, addresses the work procedure, hardwares and softwares used in this thesis. Chapter 3 also includes a presentation of the basic analysis steps used in the pipeline and different ways of performing them.

In chapter 4, *Results and Discussion*, analysis results from the processing steps presented in chapter 3 are presented and discussed using examples from the available data sets.

An overview of the produced analysis pipeline is also presented.

Chapter 5, *Conclusion*, summarizes and evaluates the work done and discusses limitations and possible future research aspects.

An analysis script and output from it is included in *Appendix*.

2

MEG Theory

MEG is a non-invasive, functional neuroimaging technique mapping the activity within the brain by measuring the magnetic fields generated by neuronal currents. A MEG system essentially consists of a helmet containing an array of sensors able to pick up on extremely weak magnetic fields, placed in close proximity to the subject's head. A state-of-the-art MEG system can be seen in Figure 2.1. To reduce the effect of environmental noise that might overwhelm the MEG signals and thereby facilitate the measurement of the weak magnetic fields of interest, MEG measurements are almost always conducted in magnetically shielded rooms.

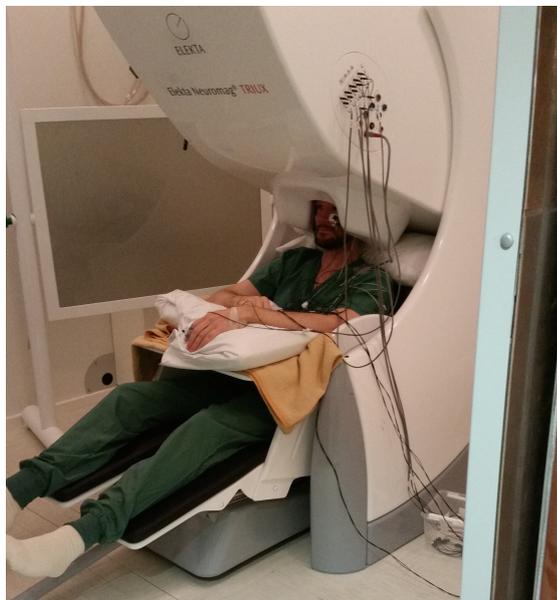


Figure 2.1: A subject in the Elekta Neuromag[®] TRIUX MEG system inside a magnetically shielded room at the NatMEG facility.

There are many challenges in MEG data analysis, one being to model the neuronal activity from the measured magnetic fields outside the head. This is the so called electromagnetic inverse problem, with one of the biggest issues being that it is underdetermined, meaning that there is no unique solution to the problem. A necessary part of solving the inverse problem is solving the more straight forward electromagnetic forward problem. The forward problem consists of computing the magnetic fields at the sensor locations from current elements using Maxwell's equations. To compute a unique solution to the inverse problem, constraints must be implemented based on prior knowledge of neuroanatomy and bioelectromagnetism such as, for example, possible source location, number of sources and their spatial extent[4].

Data from structural MRIs of the subject's brain are used for applying physiological constraints needed to solve the forward and inverse problems in MEG data processing. The MRIs are also used for visualization of the computed neuronal activity on an image of the brain, since MEG itself is unable to form structural images.

2.1 Neuroanatomy and Biomagnetism

2.1.1 Macroscopic brain

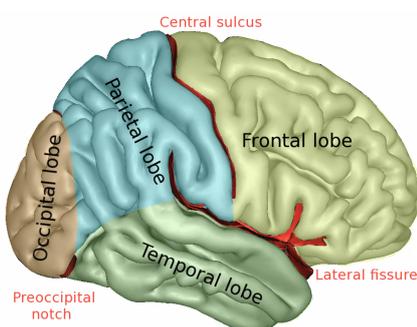


Figure 2.2: The right hemisphere with the four lobes and three prominent sulci marked. Directly posterior to the central sulcus, which marks the division of the frontal and parietal lobe, lies the primary somatosensory cortex (S1). Adapted from [11].

The largest part of the human brain called the cerebrum is divided into two hemispheres with a wrinkled surface layer called the cerebral cortex, forming deep valleys called sulci or fissures and ridges called gyri. The cerebral cortex consists of gray matter and contains the sources of signals measurable by MEG. Beneath the cerebral cortex lies the white matter in which e.g. connections between the sources are located.

Each hemisphere is divided into four regions, or lobes, according to Figure 2.2. The prominent central sulcus (also called central fissure or Rolandic fissure) marks the division of the frontal lobe and the parietal lobe. In the parietal lobe, directly posterior to the central sulcus, the postcentral gyrus rises. The postcentral gyrus is the location of the primary somatosensory cortex (often referred to as S1), the area in the human brain

connected to the sense of touch, expected to be the first brain region showing activation when a sensory stimulus is applied. The S1 region on the left hemisphere monitors the right side of the body, while the S1 region on the right hemisphere monitors the left side of the body.

2.1.2 Microscopic brain

The cerebral cortex has a thickness of a few mm and consists of at least 10^{10} neurons spread out over a total surface area of about 2500 cm^2 , folded in a complicated way creating the sulci and gyri[6]. Neurons are the information processing units of the brain and consist of a soma (or cell body), from which nerve fibres called dendrites and an axon extend. The dendrites are the neuron's input, picking up on electrochemical stimuli from other neurons and sending signals to the soma. The axon is the neuron's output with the primary function of sending electrical impulses away from the soma to other neurons, muscles or glands. The structure in which a neuron sends a signal to another

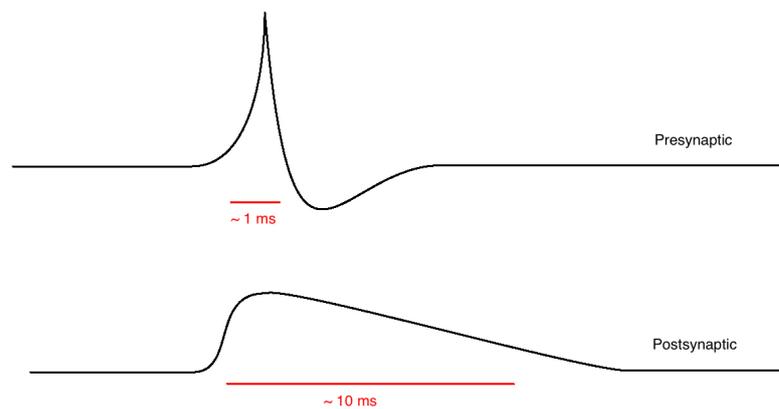


Figure 2.3: The faster, biphasic presynaptic potential and the slower, monophasic postsynaptic potential.

cell is called a synapse. The presynaptic potential, or action potential, in the axon is short lived and biphasic with a time span of approximately 1 ms[6], illustrated in Figure 2.3. The postsynaptic potential in the dendrite, also illustrated in Figure 2.3, is slow compared to the presynaptic, with a time span of 10s of ms, and it is monophasic, lacking the dip of the presynaptic potential. These potentials are orders of magnitude too small to give rise to magnetic fields measurable on the scalp, meaning that thousands of neurons must be active simultaneously to give rise to fields large enough to be picked up by MEG[6]. Presynaptic potentials don't usually contribute to fields measurable by MEG since presynaptic potentials can cancel each other out if they are not completely synchronized due to their short time spans and biphasic nature. Summation of postsynaptic potentials resulting in fields strong enough to be measured by MEG are, however,

physically feasible thanks to their longer time spans and monophasic nature.

The type of neurons mainly responsible for producing the magnetic fields measurable by MEG are called cortical pyramidal neurons, illustrated in Figure 2.4. They are named after their triangular shaped soma and have relatively long (hundreds of microns or more) dendrites that are aligned perpendicular to the surface of the cerebral cortex[7]. The alignment of the dendrites in combination with their tendency to be active simultaneously leads to the required summation of postsynaptic currents[6].

2.1.3 Neuromagnetism

According to electromagnetic theory and Maxwell’s equations electric currents give rise to magnetic fields according to the right hand rule (with the thumb pointing in the direction of the current, the fingers curl in the direction of the magnetic field), as seen in Figure 2.4. From a distance, the postsynaptic potentials look like current dipoles oriented along the dendrites of the neurons. This fact leads to the sources of the fields measured by MEG often being modelled as ideal current dipoles, which is a mathematical idealization consisting of a current element with an infinitesimal separation between the positive and negative pole and a current approaching infinity, giving rise to a finite dipole moment \mathbf{Q} , which is the product of the separation and current[12]. This means the ideal current dipole has a position, direction and magnitude but no spatial extent.

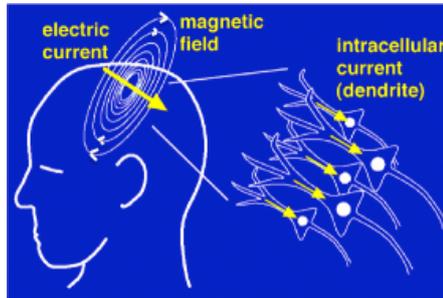


Figure 2.4: Postsynaptic currents in dendrites of aligned cortical pyramidal cells, summed up to produce magnetic fields according to the right hand rule. Adapted from [13].

The magnitude of the magnetic field from an ideal current dipole in a homogeneous space has an inverse squared dependence as a function of the distance to the source, as can be seen in Equation 2.1[14], with $\mathbf{B}(\mathbf{r})$ being the magnitude of the magnetic field at position \mathbf{r} and \mathbf{r}' being the position of current dipole.

$$\mathbf{B}(\mathbf{r}) \propto \mathbf{Q} \times \frac{\mathbf{r} - \mathbf{r}'}{\|\mathbf{r} - \mathbf{r}'\|^3} \quad (2.1)$$

This means that the magnetic field decreases drastically with the distance from the dipole. Although thousands of neurons might fire at the same time, the produced magnetic fields on the scalp level are still extremely weak, on the level of 10 – 100 fT[7]. The magnitudes of different magnetic fields we might encounter in daily life or at hospitals

are presented in Table 2.1. For example, the magnetic fields measured by MEG are six orders of magnitude smaller than the magnetic fields from urban noise. Cautions are taken due to this fact, e.g., conducting the MEG measurements in a magnetically shielded room (MSR) to avoid environmental noise sources, but the subject's body is in itself a source of artifacts, i.e. the human heart is producing magnetic fields. A common type of magnetic field sensor, fluxgate, can reach a sensitivity of around 10^{-11} [15], which is orders of magnitude greater than the fields produced by neuronal activity, emphasising the high demands on sensitive sensors for MEG measurements.

Table 2.1: Magnetic field strengths listed after order of magnitude for comparison.

Magnetic Field Strength (order of magnitude)	
MEG (measured field)	10^{-13} T [7]
Human Heart	10^{-10} T
Urban Noise	10^{-7} T [16]
Earth's Magnetic Field	10^{-5} T [17]
Refrigerator Magnet	10^{-2} T [18]
MRI (applied field)	10^0 T [17]

2.2 MEG Sensors

To be able to measure the extremely weak magnetic fields generated by neuronal activity, a combination of low noise levels and sensitive sensors is required. A common way to prominently reduce the noise level is to conduct the MEG measurements in a magnetically shielded room, making sure the subject isn't wearing any magnetic materials and using equipment adapted for MEG measurements.

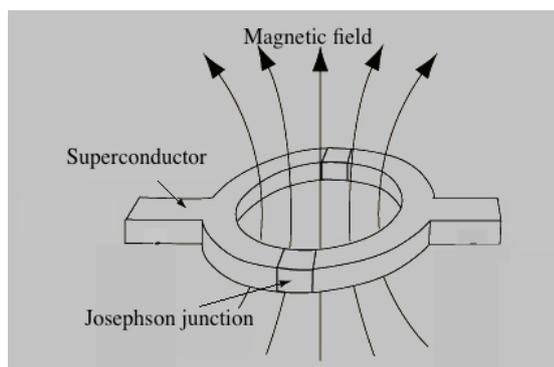


Figure 2.5: Magnetic field flowing through a SQUID, consisting of a superconducting loop interrupted by two (or one) Josephson junctions. In MEG systems, the SQUIDS are connected to pickup coil in form of magnetometers and/or gradiometers. Adapted from [19].

Today, state-of-the-art MEG systems consist of an array of hundreds of SQUID-based sensors. SQUID stands for superconducting quantum interference device and SQUIDS are extremely sensitive magnetic flux detectors, basically converting magnetic flux to voltage and can reach a sensitivity level below $1 \text{ fT}/\sqrt{\text{Hz}}$ [20]. The SQUID is, as expected by its name, based on superconductive properties and operate in a superconductive state. Certain materials become superconductive when kept at extremely low temperatures, cooled down under the material specific so called critical temperature. Liquid helium with a temperature of 4 K is used for cooling of sensors in state-of-the-art MEG systems. Examples of properties exhibited by superconductive materials are the loss of electric resistance and expulsion of magnetic flow through superconducting loops by so called screening currents emerging, counteracting the applied magnetic field. A SQUID is a superconducting loop interrupted by one or two thin insulators called Josephson junctions, over which voltage appear when a critical current over the junction is exceeded. When the SQUID is used properly, this voltage is a function of the magnetic flux through the SQUID loop. A SQUID can be seen in Figure 2.5.

Sensors optimized for MEG measurements consist of SQUIDS inductively connected to pickup coils in form of a single or multiple loops, called magnetometers and gradiometers. Magnetometers consist of a simple loop, most sensitive to currents in close proximity to but not directly underneath the loop. The two most common types of gradiometers consist of two oppositely wound loops, either in the same plane called planar gradiometers, or on the same vertical axis called axial gradiometers. Axial gradiometers have the same sensitivity pattern as magnetometers but are less sensitive to homogeneous ambient noise. This is thanks to the opposite winding of the loops: if the same flux flows through both loops, the resulting currents will have opposite directions and thereby cancel out. Planar gradiometers on the other hand have a different sensitivity pattern. They are more near-sighted than the axial gradiometers and magnetometers and are most sensitive to sources located directly underneath them. However, they have a similar noise cancellation quality as the axial gradiometers due to the opposite winding of the loops.

3

Materials and Method

3.1 Work Procedure

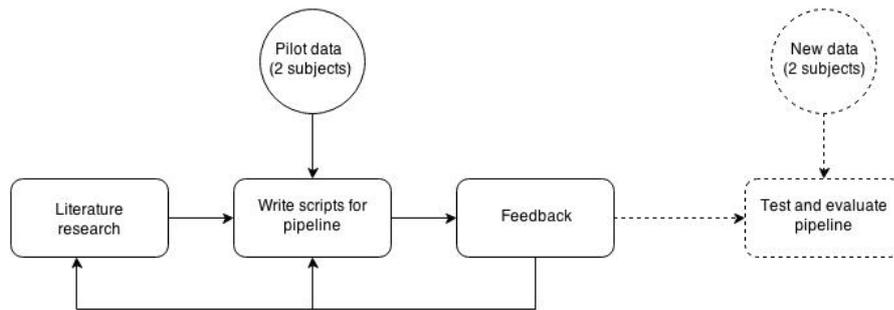


Figure 3.1: Workflow for developing the analysis pipeline, explained in detail in section 3.1.

The development of the MEG data analysis pipeline was done using an explorative, iterative method shown in Figure 3.1. By trying to analyse two raw MEG measurement data sets from pilot measurements within the Arousal Project (see Section 1.1.3), the analysis scripts were developed. The process was based on literature studies combined with continuous discussions with and feedback from technical and medical experts from MedTech West, Chalmers University of Technology, Sahlgrenska Academy and NatMEG at Karolinska Institutet. Depending on the nature of the feedback, the scriptwriting would either continue by editing existing scripts or moving forward by writing new ones, or the literature research stage would be revisited due to the realisation that a better theoretical base must be acquired before continuing the script writing. The literature studies included general theory about magnetoencephalography as well as looking at how studies using MEG are presented in scientific literature. Concepts in MEG data

analysis and ways to conduct the analysis were studied in general and combined with specific studies of the analysis in the MNE software: reading scientific papers, the MNE software manual and examples of analysis protocols in the MNE software.

After the analysis pipeline was perceived as being complete enough for further testing, two newly acquired data sets from the Arousal Project were used to test and evaluate the pipeline. These new data sets were recorded from subjects previously characterized with microneurography as being on opposite sides of the responder/non-responder spectrum. An experimental protocol corresponding well to the one used for the first two MEG pilot measurements was used when acquiring the two new data sets. Using the pipeline on new data sets was an appropriate way of testing the generality of the pipeline, since these data sets were not used to develop the pipeline and therefore more suitable for testing it.

3.2 Data Acquisition

3.2.1 Equipment

The MEG measurement data used in this thesis work was acquired at NatMEG, the swedish national facility for magnetoencephalography, with an Elekta Neuromag[®] TRIUX MEG system in a MaxShield[™] magnetically shielded room. The Elekta Neuromag[®] TRIUX system, seen in Figure 2.1, has 306 sensors in total: 102 magnetometers and 204 planar gradiometers. They are spread out over 102 locations within the MEG helmet with three overlapping sensors at each location in form of one magnetometer and two perpendicularly oriented gradiometers. The measurement data from the Neuromag system is in FIF file format, allowing effective organisation of information and favoured by the MNE software[5]. Structural MRI data for all subjects was recorded previous to this thesis work.

3.2.2 Subjects

Four healthy, male subjects were used. Recordings on two subjects were done previous to this thesis work and the other two recordings were done towards the end of this thesis work. The latter two subjects had previously been categorized as being on opposite sides of the responder/non-responder spectrum within the arousal study, using microneurography[9].

3.2.3 Experimental protocol

The experimental protocol used for the MEG measurements was similar to the protocol used for the previously done microneurography measurements[9]. An arousal stimulus in form of an electric shock just below the pain threshold was delivered to the left hand index finger with electrodes placed on the last and middle digits of the finger. The stimuli were timed with the R-wave of the QRS-complex of the ECG with two different conditions: either the stimulus was applied with no delay or with 200 ms delay with

respect to the R-wave. A total of 72 stimuli were applied, 36 for each condition, in randomised order. The interstimulus intervals, varying between 30 s, 45 s and 60 s, were also randomised to maintain the surprising effect. The same protocol was used for all subjects. The subjects were instructed to sit as still as possible, not to think about anything particular and try to avoid excessive blinking during the measurements.

3.3 MNE Software Package

The software package used for developing the analysis pipeline was the MNE software package, named after a method used for solving the electromagnetic inverse problem called minimum-norm current estimates, available in the software. MNE is an academic open-source software package for MEG and EEG data processing, consisting of three sub-packages: the original package based on compiled C code, MNE-Matlab implementing some of the MNE functionality in Matlab and MNE-Python based on Python programming language. All of the versions use the same FIF file format (compatible with the output data from Elekta Neuromag systems), enabling the user to alternate between the packages without problem. MNE-C is run using terminal commands (compatible with LINUX and Mac OSX operating systems) and includes 2 graphical user interfaces (GUIs), one for visualising and preprocessing raw measurement data and one for doing alignment between MEG and MRI data as well as visualizing and analyzing the data on source level, i.e. looking at the computed activity within the brain rather than the measured magnetic fields outside. MNE-Python takes advantage of core Python libraries like NumPy, SciPy and matplotlib for scientific computations and visualizations.

Within this project, the focus has been on using MNE-Python for most of the analysis with some help of core functions and GUIs from C-based MNE. The main reason for choosing Python as the basis of scripting was the fact that MNE-Python is under constant development with continuous implementation of new functions, having a peer review process among developers to ensure a high implementation level and theoretical basis for new code[5]. Python is also a user-friendly programming language making the scripts easy to read. There are many benefits with using a scripting based analysis pipeline, as can be done in MNE-Python, instead of manually analyzing the measurement data in GUIs. For example, when analysing data for many subjects, a scripting based pipeline is much quicker and it also ensures that each data set is processed in the same way, facilitating reproducibility of results. However, some functionality is still not implemented in MNE-Python and there are benefits with using GUIs for e.g. raw data inspection, leading to an analysis pipeline not solely based in Python.

Other softwares were also used to enable the data analysis in MNE. During this thesis work, Canopy[21] was used as the Python analysis environment, and to perform complete data analysis MNE was used in combination with FreeSurfer[22], which is an open source software for processing MRI data and vital for getting the MEG data analysis to source level.

3.4 Developing the Analysis Pipeline

The analysis pipeline was developed by having an analysis aim to compare and possibly find differences between the two different stimuli conditions in the data sets - either the stimulus was applied on the R-wave of the ECG or 200 ms after. The analysis aim was set to have an analysis purpose when developing the pipeline and is not the primary analysis aim within the arousal project. The focus has mainly been on the preprocessing of measurement data and analysis on sensor level (studying the magnetic fields rather than the computed neuronal activity). This was due to the fact that mistakes or lack of for example sufficient noise reduction in data processing on sensor level will propagate to source level, and in a worst case scenario this could lead to source estimates completely unrelated to the actual neuronal activity. Another reason for focusing on sensor level is the basic fact that if activity cannot be seen on sensor level, it will not appear on source level either.

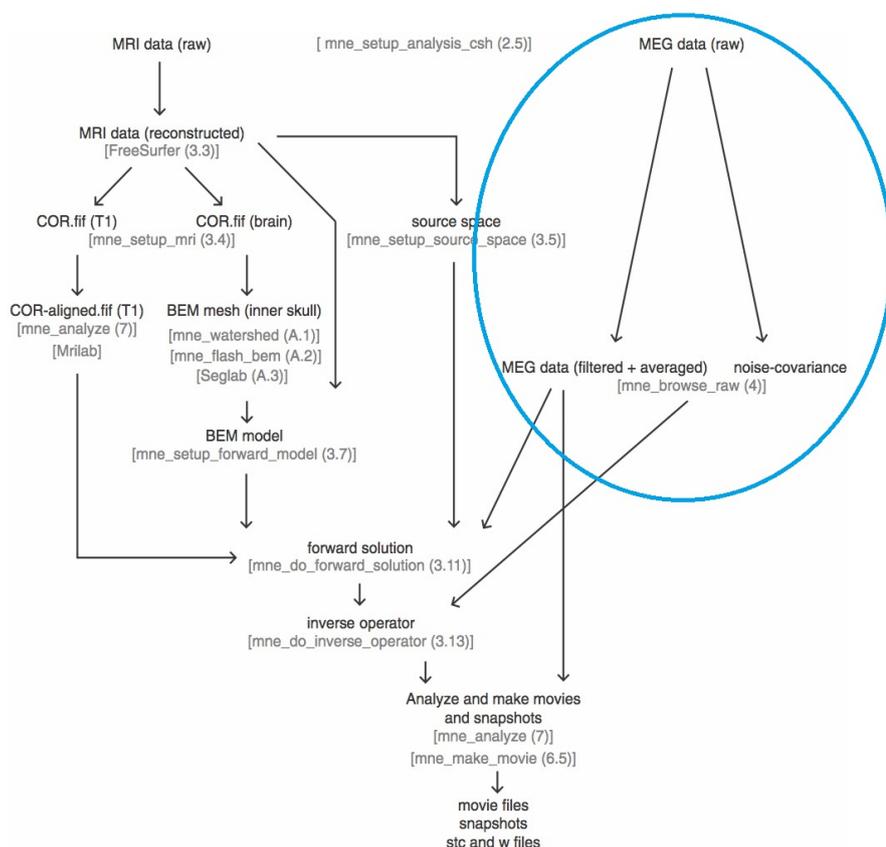


Figure 3.2: MNE-C workflow from the MNE software User’s Guide. The focus of this thesis work has been the processing of pure MEG data, which is circled. The references in parenthesis refers to sections in the MNE software User’s Guide. Figure used and altered with permission from [23].

The essential MEG data processing steps and methods for implementing them are presented in this section. Figure 3.2 shows the MNE processing/analysis flow chart from the MNE-C manual[23], which gives a good overview of how MEG data analysis can be done. In this thesis work, the focus has been on developing an appropriate processing pipeline for the MEG data, corresponding to the circled part of the figure, rather than the processing of MRI data, which is a necessary part for analysing the MEG data on source level.

3.4.1 Preprocessing

The first step of the analysis is to visually inspect the raw measurement data. It is important to make sure there are no bad sensors (referred to as channels hereafter), which are easy to spot visually since they are either completely flat or extremely noisy, as can be seen in Figure 3.3. During visual inspection data segments including large artifacts that might corrupt the data during further processing are also sought after. The effect of eye blinks on measurement data is studied in this first step, by comparing activity in the EOG channels with regular MEG channels. If bad channels or data segments are found, they are to be marked and removed. The visual inspection was done in the MNE-C GUI for raw data inspection called *mne_browse_raw*, in which bad channels can be marked and user defined events can be saved for marking bad data segments.

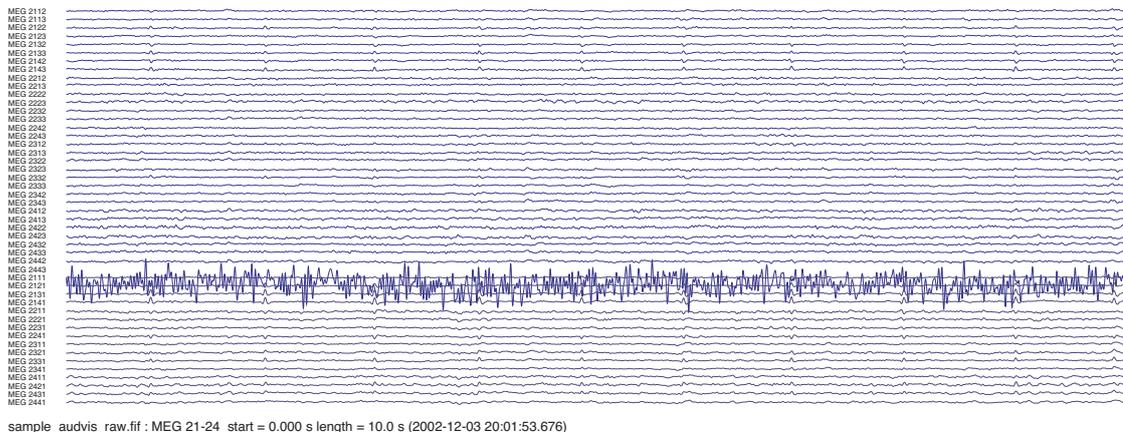


Figure 3.3: A snapshot from MNE-C GUI *mne_browse_raw* showing a 10 s segment of measurement data from a selection of MEG channels listed to the left, with one bad channel clearly visible. The data set shown is from the MNE software sample data set, included in the software.

After the visual inspection and marking of possible bad channels and segments, the measurement data was downsampled to 300 Hz for faster processing. Thereafter, the power spectral density (PSD) plot was inspected and the data was band-pass between 0.6 – 45 Hz and notch filtered for 50, 100, 150Hz. The band-pass filter cutoff frequencies were chosen in a similar fashion as example scripts from the MNE webpage and work-

shops at the NatMEG facility. The notch filter was applied to make sure the power line frequency and multiples of it was removed. Inspection of the PSD plot of the filtered data was done to make sure the filtering worked properly.

3.4.2 Averaging

Since the objective of the analysis is to investigate how the subjects respond to stressful stimuli, the interesting segments of the measurement data are the time intervals around each stimulus event. Rather than looking at the complete set of continuous measurement data, the data was cut into user defined time spans around the stimulus events in a process called epoching. This was done with the help of a channel in the MEG measurement data containing the timing of each stimuli. Thereafter the epochs were averaged to form an evoked signal, with the benefit of increasing the signal to noise ratio due to random noise being averaged out. An evoked signal was created for all 72 epochs as well as two separate evoked signals containing the two conditions separately. To separate the two conditions, an excel sheet containing the stimulus protocol was read.

Epoching was also done based on the R-wave of the ECG to create an evoked signal of the cardiac artifact. The ECG-epochs coinciding with a stimulus event were not used for creating the ECG-evoked signal.

3.4.3 Artifact and Noise Reduction

The magnetic fields measured by MEG are extremely weak and can easily be affected or even overwhelmed by artifacts and noise stemming from other sources than the brain. Heartbeats, eye blinks and muscle activity are the most prominent sources of biological artifacts[24], and will unavoidably affect the measurement data. The removal of cardiac artifacts is particularly important for the analysis pipeline developed for the Arousal Project, due to the fact that the stimuli are time locked with the heartbeat, applied either at the R-wave or 200 ms after. This means that the artifact caused by the heartbeat will be prominent in all trials and won't average out as it might do in other experimental protocols. Specifically due to the analysis aim used for the development of the analysis pipeline, which is comparing the two stimulus conditions, the importance of removing or at least reducing the cardiac artifact significantly is extremely important, since a discovered difference between the two conditions might actually only be caused by the relative timing between the heartbeats and the stimuli.

Table 3.1: How to handle different types of noise and artifacts in MEG measurement data.

Artifact	Method
Small, randomly distributed (noise)	Reduced by averaging
Large, uncommon (bad data segments)	Manually removed from analysis
Large, recurring (artifacts)	Reduced by signal processing methods

For a quick overview of how different types of artifacts in MEG data are handled, see Table 3.1. Randomly distributed, low-amplitude artifacts do not usually give rise to big problems since they tend to become sufficiently reduced when averaging the data. One common strategy for handling large-amplitude artifacts is to simply remove the data segment or trial containing the artifact from further analysis. However, this is a feasible approach only if the artifact is rare - removing a large fraction of trials will jeopardize the reliability of the analysis results. For artifacts recurring in a large portion of trials, signal processing methods have been developed for reliably reducing their effect.

At the NetMEG facility, an Elekta software called MaxFilter can be used to apply either regular or temporal signal space separation (SSS or tSSS) to the measurement data. Two other methods for artifact reduction are available in MNE: signal space projection (SSP) and independent component analysis (ICA). The different methods were tested and compared based on feedback from coworkers and experts and visual inspection of their results on the two pilot measurement data sets.

Signal Space Separation

Signal space separation (SSS) is a mathematical method based on separating sources outside the MEG sensor array from sources inside, with the hope of removing all environmental noise sources as well as biological artifacts stemming from locations outside the MEG sensor array while keeping the signals from the brain. The method utilizes knowledge about the fundamental physical properties of electromagnetic fields and the sensor configuration geometry, with the only assumptions made being that all sensors are located in a current free volume with the sources of magnetic fields located at least a few cm away from the sensors[25].

In an N-channel measurement system, the signal space is an N-dimensional space. Due to the fact that the sensors are located in a volume free from magnetic sources, harmonic function expansions of the signal space can be used as a basis in which the measured signal (the magnetic field) can be uniquely represented. Simply put, there are separate basis vectors for signals stemming from inside and outside the sensory array and the signal space can thereby be divided into two subspaces - signal space separation. The subspace corresponding to signals from sources outside the sensor array can be left out, keeping only the signals from sources inside it.

SSS is a purely spatial filtering method, whereas temporal SSS (tSSS) also accounts for the time dependency. Like SSS, tSSS removes noise and artifacts produced by sources outside the sensor array by dividing the sensor space into two subspaces, but tSSS is also able to remove artifacts produced by sources in close proximity to the array, like noise from the sensors themselves or noise sources on the scalp of the subject. Due to the assumption of the immediate vicinity of the sensors being free from magnetic sources, possible interference sources in this volume will be partly represented in the subspace outside the sensor array and partly in the subspace within the array. What tSSS does is to extract the components in the two subspaces showing a strong temporal correlation with each other, which shouldn't affect the signals coming from the brain since they are temporally uncorrelated with the interference signals[26].

In this thesis work, regular SSS was applied to all four data sets prior to analysis with the MNE software. For comparison reasons tSSS with three different correlation thresholds of 0.98 (the default value), 0.90 and 0.80, were applied to one of the pilot data sets. This was done by staff at the NatMEG facility using the Elekta Neuromag MaxFilter software.

Signal Space Projection

Signal space projection (SSP) is a method using statistical signal analysis, based on the assumption that the environmental noise and artifacts are caused by reasonably stable statistical processes[4].

In the N-dimensional signal space, different configurations of the magnetic field (corresponding to different sources) are represented by different vectors[27]. The basis of the SSP method is finding M vectors representing the magnetic fields from sources of noise or artifacts like the heartbeat, and then dividing the signal space into two subspaces: one parallel and one orthogonal to these vectors. This is done by projecting the N-dimensional signal space to an "N minus M"-dimensional space orthogonal to the M principal components of the artifact or noise, and thereby removing them from the signal space. For this method to produce a reasonable estimate of the brain signals space, the M-dimensional noise space must be nearly orthogonal to and of a much lower dimension than the brain signal space. Otherwise the orthogonal projection will distort the signals, which in further data analysis will give false current estimates.

In this thesis work, SSP was tested for cardiac artifact removal by using the inbuilt function in the MNE software with default parameters.

Independent Component Analysis

Independent component analysis (ICA) is a method for separating noise and artifacts from the rest of the measurement data, by dividing the data into signals representing different sources and removing the ones representing noise and artifacts. The division is based on the assumption that these signals are statistically independent and nongaussian[24][28], and the measured data is at each time point a mixture of these statistically independent signals, referred to as independent components (ICs). The difficulty lies within the fact that neither the source signals nor their mixing process is known, all that is known is the measured mixed signal. To solve this problem, it is assumed that the mixing of the ICs is a linear process and the number of ICs can never exceed the number of sensors[24].

The separation of ICs is implemented differently in different ICA algorithms, but is built on optimizing a contrast function, which can be based on for example entropy, mutual independence or high-order decorrelations[4]. The infomax principle and the fastICA algorithm are two examples of implementation of ICA, based on a contrast function which maximizes the output entropy and an iterative method based maximization of negentropy, respectively. For further information about ICA algorithms, see Hyvärinen et al.[28].

When the ICs have been identified, the next step is to extract the ones related to the artifacts and thereby separating the brain signals from unwanted interfering signals. In the MNE software, the components can be compared with data segments corresponding to known artifacts (extracted with the help of electrocardiography, ECG, and electrooculography, EOG, channels) either by using regular Pearson correlation or cross trial phase statistics (CTPS), available for detection of artifacts related to heartbeats. CTPS uses phase histograms calculated across epochs centered around the R-wave of the heartbeat. Magnetic fields unrelated to the cardiac activity will have a uniform CTPS, whereas the magnetic fields produced by cardiac activity are synchronous with the heartbeat, producing a nonuniform phase distribution[24]. CTPS is recommended in literature and the default method in the MNE software for finding components related to cardiac artifacts, due to the fact that some of the components related to cardiac artifacts have weak amplitudes, making them difficult to detect with correlation in the amplitude domain[24].

The MNE software has inbuilt semiautomatic functions for detection and removal of artifacts related to cardiac activity and eye blinks using ICA. Three different ICA methods are available: InfoMax, Extended InfoMax and FastICA, with FastICA being the default alternative. For choosing components related to eye blink artifacts, Pearson correlation is the only available method in the MNE software, while the alternative of choosing CTPS is available for detection of cardiac artifact components. When using CTPS, the MNE software provides scores of the ICs based on Kuiper's test of uniform phase distribution[24], with higher scores indicating nonuniform distribution and thereby indicating a relation to the heartbeat.

All three ICA methods and the two methods for picking components were tested for cardiac artifacts. For eye blink artifacts FastICA and Pearson correlation was used. Visual inspection of the components with high scores were used for final decision of components to be removed.

3.4.4 Statistical analysis

The impact of the cardiac artifact on the measurement data was investigated by statistical methods. The analysis was done on sensor level, channel by channel, comparing the dependence levels for each condition versus the cardiac artifact using a two-tailed t-test. The cardiac artifact data used in the t-test were 36 randomly selected ECG-epochs (the same number of epochs as for each stimulus condition) not coinciding with the stimulus epochs. The two conditions were also compared to each other with a t-test to investigate possible differences between them, which was the analysis aim. This was done on preprocessed data without cardiac artifact removal as well as on data in which ICA (FastICA with CTPS) was used for cardiac artifact removal.

3.4.5 Forward Solution

To be able to construct the neuronal currents linked to stimulus processing from the sensor level data, that is, to solve the inverse problem, the forward problem must first be solved. A forward solution in MEG analysis is a computation of the magnetic fields

at the sensor locations based on the neuronal activity. To be able to do these forward computations, the magnetic properties of the head, the type and location of sources and sensors including sensor orientations and pick-up loop geometries are needed, requiring a number of processing steps.

The current dipole is the elementary source model used in MNE and the set of possible locations of dipoles form the so called source space. The source space used in this thesis work consists of 10 242 source locations in each hemisphere, giving a spacing of approximately 3.1 mm for the average brain. The magnetic permeability of the head was assumed to be constant over scalp, skull and brain, and the same as for free space[29], leading to the choice of a single compartment boundary element model for computing the head model using the FreeSurfer software package. Since MNE supports the sensor properties for Elekta Neuromag systems, no additional sensor information had to be manually added. To compute the forward operator, the two different coordinate systems stemming from the structural MRI data and the MEG measurement data respectively must be aligned. This processing step is called co-registration and must be done manually by the user, connecting the fiducial landmarks (nasion and pre-auricular) from the MEG data to the MRI head model. This was done in MNE-C GUI *mne_analyze*.

3.4.6 Inverse Solution

Source reconstruction was done by computing an inverse operator and then applying it to measurement data. The MNE software implements linear inverse methods based on minimum-norm estimates. Along with the forward operator, a noise covariance matrix is also needed for computation of the inverse operator. The noise covariance matrix was computed based on the time span (-0.175 , -0.25) s before each stimulus event using an inbuilt function with default settings. This specific time span was chosen to avoid possible effects of the cardiac artifact centered around -0.2 s and 0 s respectively for the two conditions.

Dynamic statistical parameter mapping (dSPM), based on minimal-norm estimates and transforming the reconstructed current density values into dimensionless statistical quantities, was the main method used for analyzing source level data. Aside from estimating the current densities, equivalent current dipoles can easily be estimated using the MNE-C GUI *mne_analyze*. The dipole fitting is limited by the fact that only a single dipole can be fitted at one time point.

3.4.7 Script Editing

After the different processing steps were done, the analysis scripts were revised, making sure they could be executed effectively by saving processed data after each time demanding processing step, so that it could be loaded in the next step instead of having to redo time demanding processes. Another aim was to make the analysis scripts sufficiently user friendly by commenting them in a concise but descriptive way throughout all processing steps. A concise user manual was also written, describing the steps in the analysis pipeline including the order in which processing scripts should be run, parameters that

might be of interest to change and things for the user to pay attention to in the outputs from the scripts.

3.4.8 Testing on New Data Sets

For the testing of the analysis pipeline a single script for complete sensor level processing was also produced in addition to the separate processing scripts made for each processing step. The sensor level script contained downsampling, filtering, averaging, cardiac and eye blink artifact removal with ICA and was applied after visual inspection of raw MEG data and marking of bad channels and data segments. The reason for having one script doing multiple processing steps was to get a quick overview of the data set, which was facilitated by including the MNE Report function, creating a report containing user specified images, in this case plots of power spectral density, the evoked signal on sensor level and images related to the artifact removal with ICA. The sensor level analysis script was tested on two newly acquired data sets.

4

Results and Discussion

In Section 4.1 the processing steps presented in Section 3.4 are further investigated within the context of the analysis pipeline under development and data from the Arousal Project. Examples of analysis results from the available data sets are also presented. A concise overview of the developed pipeline can be found in Section 4.3. For further insight into the analysis pipeline and analysis results see Appendix A, where the complete sensor level analysis script and generated report can be found.

4.1 Analysis results

4.1.1 Preprocessing

The MNE-C GUI *mne_browse_raw* provides easy raw data visualization, enabling the user to pick channels and time spans to view. No bad channels were found in any of the four data sets, but some bad data segments were detected, as can be seen in Figure 4.1. It is difficult to determine the exact cause of these artifacts, but one possible reason is sudden movements like coughing or sneezing. The cause of the artifact, however, is unimportant in the sense of data analysis, where the main objective is to ensure that the affected data segments do not compromise the analysis. The affected data segments were therefore marked and excluded from further analysis.

Artifacts caused by eye blinks were also detected during visual inspection of the raw data, mainly in the frontal channels which is expected, as they are located in close proximity to the eyes. In the analyzed data sets, the eye blink artifacts appeared around 100 ms after essentially each stimulus event, lasting for approximately 100 ms, as can be seen in Figure 4.2. This is an expected involuntary response to arousal stimulus and can not be avoided when using this particular experimental protocol. It is therefore meaningless to remove all epochs containing these artifacts, since that would mean removing essentially all data segments containing stimuli responses. For this experimental protocol, focus should instead be on decreasing the effect of the artifacts by signal processing methods.

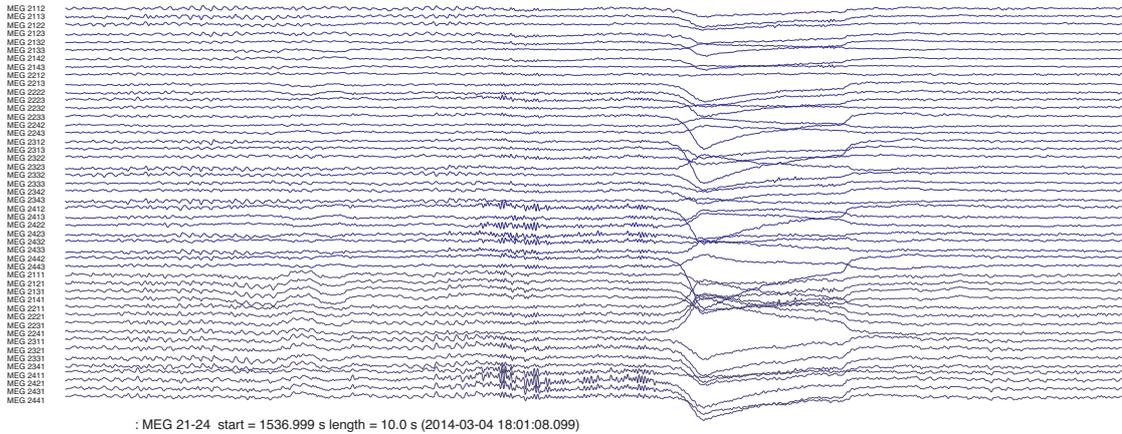


Figure 4.1: A snapshot from MNE-C GUI *mne_browse_raw* showing a 10s time window of raw measurement data containing a bad data segment with deflections up to approximately 10 pT for magnetometers. From measurements on subject 1.

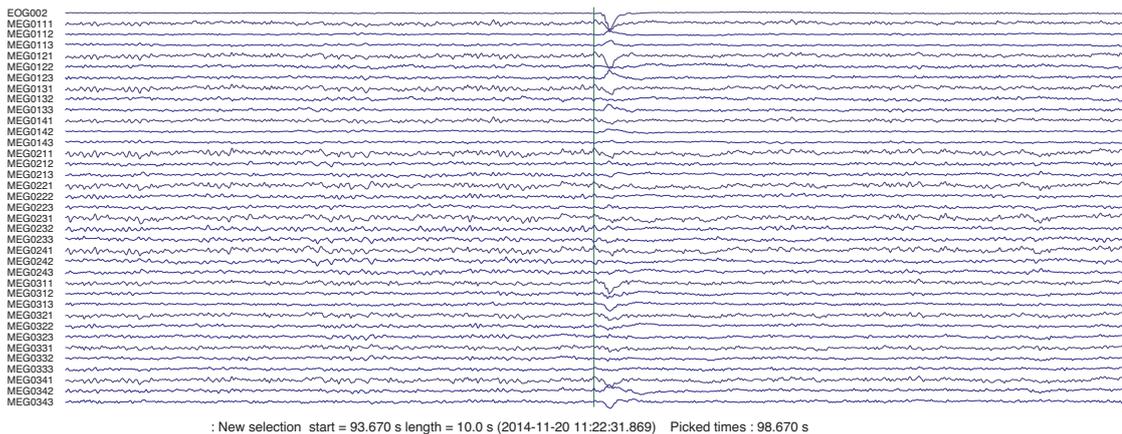
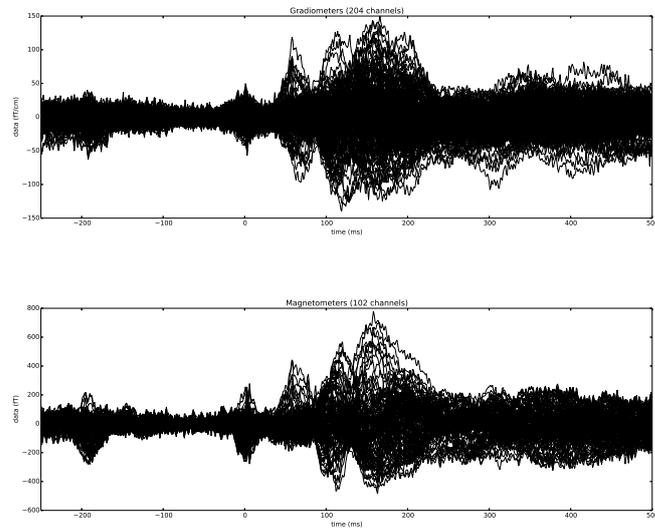
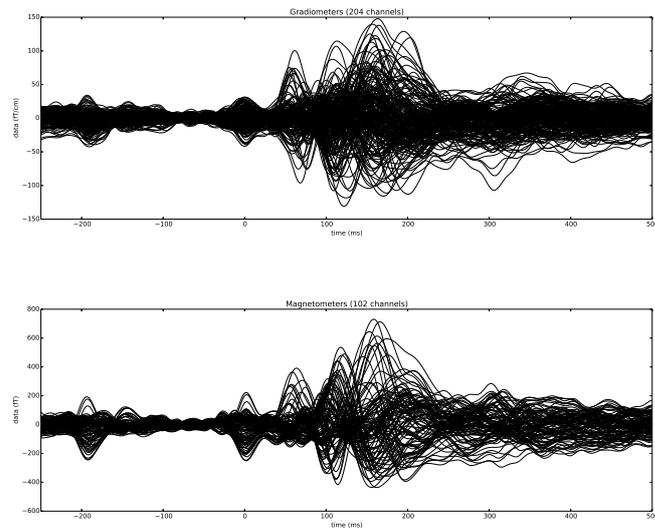


Figure 4.2: A snapshot from MNE-C GUI *mne_browse_raw* showing a 10s data segment containing an eye blink artifact with deflections up to approximately 2.8 pT for magnetometers. Except for the top one, which is an EOG channel, the channels listed to the left are MEG sensors covering the frontal parts of the subject's head. The vertical line indicates the timing of the stimulus. From measurements on subject 4.

The data sets were differently affected by noise, i.e. the measurements on Subject 1 were strongly affected by 100 Hz noise, most probably stemming from insufficient shielding of power cords. This was strongly reduced by band-pass filtering the data between 0.6 – 45 Hz combined with notch filtering of multiples of 50 Hz, as can be seen in Figure 4.3. The filtering did not visually affect the shape of the evoked signal in the time domain except for removing the 100 Hz noise, but if frequency analysis is to be done in the future, the filtering needs to be reconsidered.

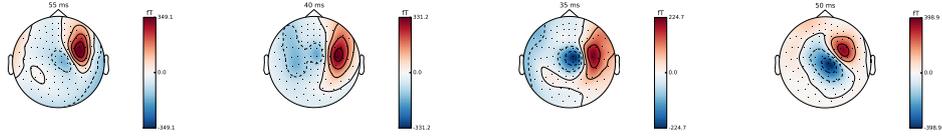


(a) Raw evoked signal averaged over 72 epochs with the stimulus at time 0.



(b) Evoked signal from filtered data (0.6-45 Hz band-pass and 50 Hz notch filter) averaged over 72 epochs with the stimulus at time 0.

Figure 4.3: Evoked data for subject 1 before (4.3(a)) and after (4.3(b)) filtering, with the main aim to reduce the 100 Hz noise. Corresponding power spectral density plots are included in Appendix B.



(a) Subject 1, 55ms after stimulus. (b) Subject 2, 40ms after stimulus. (c) Subject 3, 35ms after stimulus. (d) Subject 4, 50ms after stimulus.

Figure 4.4: The first strong peak of the evoked signal for all four subjects, located around the area of the primary somatosensory cortex of the right hemisphere. Red areas indicate magnetic fields directed out of the image and blue areas indicate fields into the image. According to the right hand rule, the dipole causing the field is located between the red and blue maximum pointing towards the back of the subject’s head. The subject specific timing of the peak is most likely related to the height of the subject, closely linked to the the distance the nerve signal has to travel before reaching the brain.

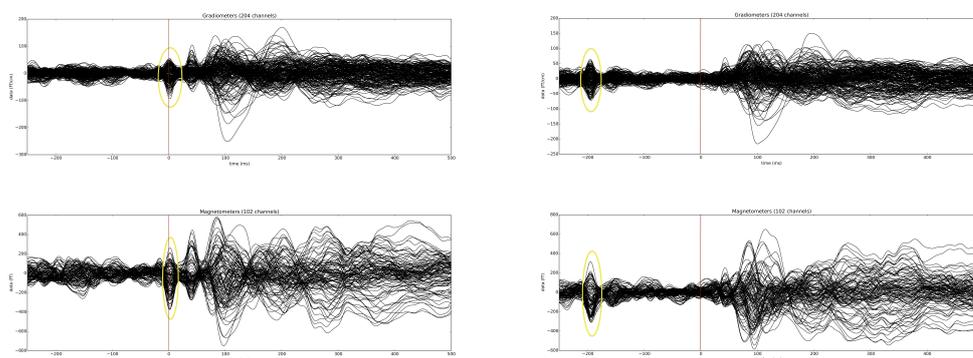
4.1.2 Averaging and Artifact Removal

A clear evoked response to the stimuli can be seen on sensor level when inspecting the averaged signals, as seen in Figure 4.3(b) for example. The first distinct peak on sensor level data for all four subjects appeared around 35 – 55 ms after each stimulus and was located around the area of the primary somatosensory cortex (S1), as shown in Figure 4.4. An activation in S1 was expected to be seen early in the processing since the response is to a sensory stimulus. The time difference between the subjects is most likely caused by the different heights of the subjects (subject 1 and 4 were definitely taller than subject 2 and 3). This leads to differences in distances between the index finger (where the stimulus is applied) and brain (where the reaction to the stimulus is measured by MEG) for the nerve signal to propagate, resulting in small time differences.

In these data sets, there is however one problem which easily can be spotted in the evoked response: the cardiac artifact is present, clearly visible at 0 ms for the first condition (stimuli applied at the R-wave of the ECG) and at -200 ms for the second condition (stimuli applied 200 ms after the R-wave), as seen in Figure 4.5(a) and 4.5(b) respectively, showing the conditions separately. The evoked response based on the heartbeat (over 3000 averaged epochs) is shown in Figure 4.6.

Signal Space Separation

Data without SSS applied was unavailable during this project, making the effect SSS has on reduction of environmental noise and physiological artifacts difficult to analyse. What can be stated is that SSS is unable to remove the cardiac artifact since it is visible in the measurement data, as previously seen in Figure 4.5. However, it is likely that the artifact is decreased by SSS since it’s amplitude is in the same order of magnitude as the stimulus evoked signal, whereas biological artifacts several orders of magnitude higher than evoked signals are not uncommon[24]. Temporal SSS with a correlation score of



(a) Evoked signal for condition 1 (stimulus timed with the R-wave of the ECG), averaged over 36 epochs.

(b) Evoked signal for condition 2 (stimulus applied 200 ms after the R-wave of the ECG), averaged over 36 epochs.

Figure 4.5: The cardiac artifact is visible for both conditions as seen in 4.5(a) and 4.5(b), where the stimulus is indicated by a vertical line and the artifact is circled. Plots for subject 2.

0.98 showed no visual difference to regular SSS, but tSSS with a correlation score of 0.80 reduced the cardiac artifact as compared to SSS, as seen in Figure 4.6(a) and 4.6(b), without affecting the rest of the stimulus evoked data visibly as compared to regular SSS.

There are two possible explanations to why SSS is not able to remove the cardiac artifact, which it theoretically should be able to do since the source of the artifact (the beating heart) is located far outside the sensor array. Either SSS doesn't work as well in practice as it does in theory or some part of the source of the cardiac artifact is located inside the sensor array, with the most reasonable explanation being a combination of both alternatives. The SSS method isn't perfect, for example, the harmonic function expansions in SSS consists of infinite sums, which is impossible to do in practical calculations. Although the subjects' hearts are located outside the sensor array, there may be parts of the cardiac artifact coming from sources within the array. It could be brain processing, blood pressure changes and motion artifacts related to the heartbeat or currents/fields propagating through the body from the heart to the head. The fact that tSSS is able to reduce the cardiac artifact further than regular SSS indicates that there are heartbeat related sources inside the sensor array temporally correlated with sources outside.

Signal Space Projection

The results obtained when using SSP for cardiac artifact reduction can be seen in Figure 4.6(c), showing the effect on the ECG-evoked data. SSP is able to decrease the cardiac artifact, but it affects the stimulus evoked data at the same time, significantly reducing

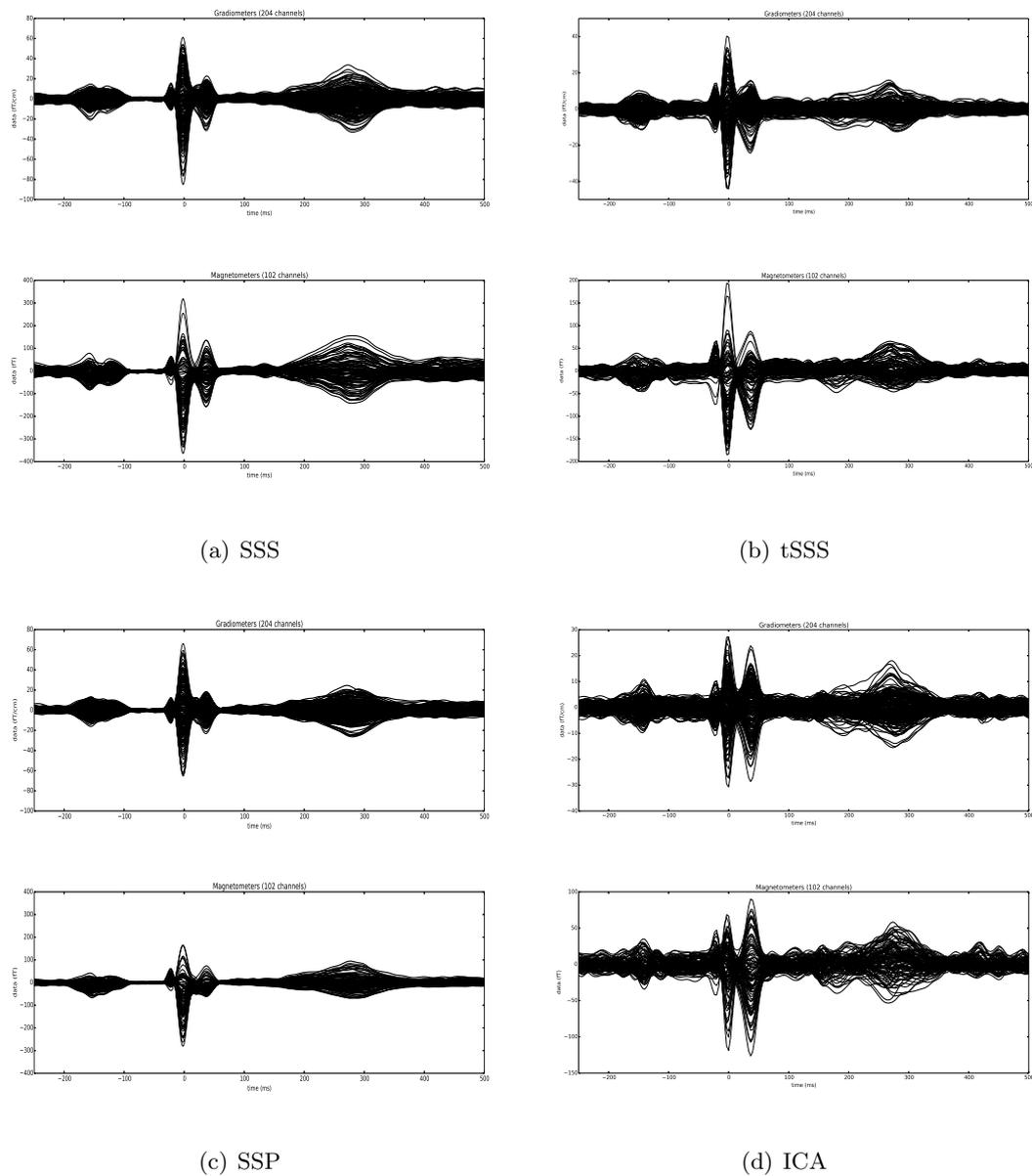


Figure 4.6: The ECG-evoked signal representing the cardiac artifact in the data. 4.6(a) shows the SSS processed signal, 4.6(b) the tSSS processed signal, 4.6(c) the SSS processed signal using SSP for artifact reduction and 4.6(d) the SSS processed signal using ICA for artifact removal. Notice that the scales on the Y-axis differ in the different plots. All for subject 2.

it's amplitude (by approximately 25%). The cause of this might partly be the method not being able to accurately estimate the artifact subspace, partly the brain signal subspace and artifact subspace not being completely orthogonal. Another factor affecting the SSP results is the difficulty in choosing projections to apply, which in the MNE software is done manually by visual inspection.

Independent Component Analysis

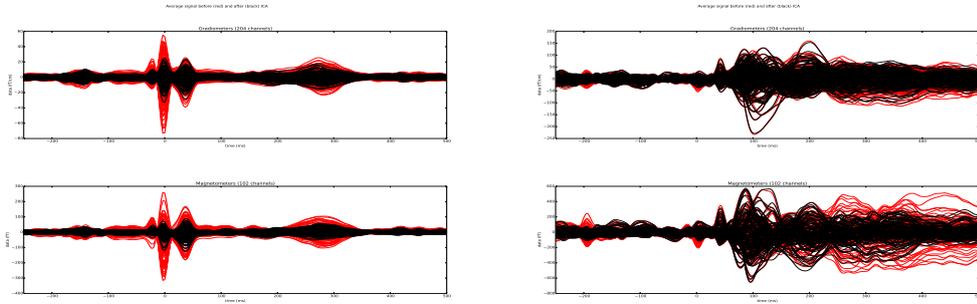
All three ICA methods available in the MNE software (FastICA, Infomax and Extended-Infomax) were tested for cardiac artifact removal on the two pilot measurement data sets (subject 1 and subject 2), yielding almost identical results. This in combination with the fact that ICA methods are theoretically complicated and by themselves worthy of a PhD thesis led to the pragmatic approach of using the FastICA method, which is the default method in the MNE software.

CTPS outperformed Pearson correlation in terms of finding ICs related to the cardiac artifact: where CTPS found multiple ICs, Pearson correlation didn't match a single IC to the cardiac artifact for the two pilot measurement data sets, resulting in an absence of artifact removal.

The visualization options for ICA in the MNE software are very good. For example, the calculated independent components with the highest artifact matching scores (based on either correlation or CTPS) can be easily plotted and overlay plots illustrating how the artifact removal affects the data is also available, as can be seen in Figure 4.7, showing the ECG-evoked signal and the the stimulus evoked signal respectively for cardiac artifact removal for. Although the cardiac artifact wasn't completely removed by using ICA for this subject, it was significantly reduced while the remaining measurement data was relatively unaffected. However, the strength of the ECG artifact differed quite a lot between the subjects, peaking at 300 fT (magnetometers) for some subjects and at 700 fT (magnetometers) for other, resulting in varying artifact removal results for the different subjects. The removal of ICs related to cardiac artifacts led to an amplitude decrease of 50 – 80% of the the ECG-evoked signal and an amplitude change of approximately 0 – 25% for the highest activation of the stimulus evoked signal, depending on subject.

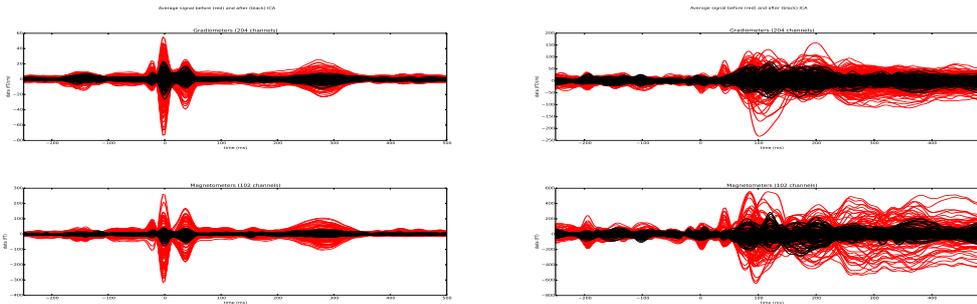
ICA was also used for removing the eye blink artifact, using Pearson correlation and the EOG-channels for finding components representing the artifacts in the data. Topographical maps of the chosen components can be seen in Figure 4.8, in which the components visually appear to be related to eye blinks or eye movements, with the fields appearing in sensors located close to the eyes. For more plots related to artifact removal with ICA, see Appendix B.

The choice of ICs to be removed from the data is not always straight forward. It can be tempting to remove more components than automatically chosen by the inbuilt functions to further decrease the artifact. It is then important to study how the removal of ICs affect the stimulus evoked signal and not only look at the artifact separately. In figure 4.7 a comparison of the same data set with different amounts of ICs removed are shown. However, in the MNE software, the visualization options in combination with inbuilt functions for matching ICs to artifacts simplifies process greatly!



(a) ECG-evoked signal averaged over more than 3000 epochs, R-wave at time = 0 ms. The removed ICs were automatically chosen by using FastICA in combination with CTPS.

(b) Evoked signal averaged over 72 epochs, stimulus applied at time = 0 ms. The removed ICs were automatically chosen by using FastICA in combination with CTPS.



(c) ECG-evoked signal averaged over more than 3000 epochs, R-wave at time = 0 ms. The removed ICs were automatically chosen by using FastICA in combination with CTPS combined with manually added ICs for a stronger reduction of the cardiac artifact.

(d) Evoked signal averaged over 72 epochs, stimulus applied at time = 0 ms. The removed ICs were automatically chosen by using FastICA in combination with CTPS combined with manually added ICs for a stronger reduction of the cardiac artifact.

Figure 4.7: Overlay plots showing how ICA used for cardiac artifact removal affects the measurement data for ECG-evoked signal as well as stimulus evoked signal. Red lines indicate signal before artifact removal and black lines after. In 4.7(a) and 4.7(b) the effect of removing the automatically chosen components can be seen, and in 4.7(c) and 4.7(d) the effect of manually removing more components for further artifact reduction can be seen. Although manually increasing the amount of removed components reduces the magnitude of the artifact itself, as can be seen in 4.7(c), the stimulus evoked signal is affected in a way that suggests important information might be lost, 4.7(d). Plots for subject 2.

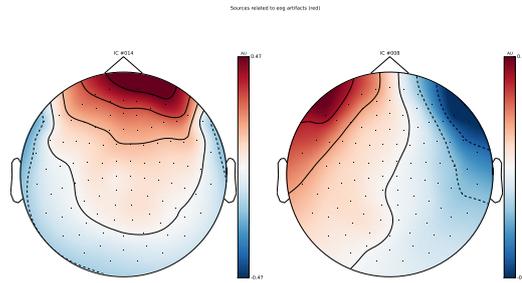


Figure 4.8: Topographical view of two ICs related to eye blink artifacts using ICA.

Comparison of Artifact Removal Methods

The largest reduction of the cardiac artifact was done by ICA in combination with regular SSS, decreasing the amplitude of the ECG-evoked signal by approximately 70% for subject 2. TSSS with a correlation score of 0.80 and SSP combined with SSS were also able to decrease the cardiac artifact, but not to as high a degree as ICA. Furthermore, SSP strongly affected the stimulus evoked signal, whereas ICA and tSSS only affected it marginally. A combination of tSSS and ICA was also tested, but gave similar results as SSS combined with ICA. The fact that the two pilot measurement data sets were SSS processed and that the Elekta software needed for applying tSSS were only accessible from NatMEG led to the decision of using SSS processed data in the pipeline, and combining it with ICA for physiological artifact removal. Not only did ICA provide a stronger reduction of the cardiac artifact than SSP, ICA is also a more user friendly method for artifact removal than SSP in the MNE software due to the more advanced visualization alternatives, facilitating the picking of components and overviewing of results. Choosing ICA over SSP for physiological artifact removal is also supported in literature, where Gramfort et al. recommends using SSP for removal of environmental noise and then ICA for physiological artifacts[29]. Since SSS is used for environmental noise reduction in this project, SSP is left out of the analysis process.

4.1.3 Statistical Analysis

At certain time spans statistical differences between the two conditions can be seen at the same time as there is a statistical independence of the stimulus evoked and ECG-evoked signals. This is shown in Figure 4.9 for a channel located close to S1 in the right hemisphere, with the time of interest being around 40 ms after stimulus, which is the time of the S1 peak for subject 2, previously shown in Figure 4.4(b). The question is whether this difference between the two conditions is caused by the time shift of the cardiac artifact, represented by the ECG-evoked signal in Figure 4.9, or whether the difference is independent of the cardiac artifact. Although the stimulus evoked signals are statistically different from the ECG-evoked signal for large time spans, it doesn't

mean that they are unaffected by it. The stimulus evoked signal contains the cardiac artifact, and can be thought of as a combination of a "clean" stimulus evoked signal and the ECG-evoked signal. An amplitude difference of the time shifted ECG-evoked signals is clearly visible around the time span of interest, with a higher amplitude 40 ms after after the R-wave (time of interest for condition 1) than 240 ms after the R-wave (time of interest for condition 2), as can be seen when comparing the upper two plots in Figure 4.9. The same statistical test done on data on which ICA was applied for cardiac artifact removal can be seen in Figure 4.10, and it is evident that the time span for which statistical independence between the two conditions can be seen is strongly decreased as compared to data without artifact removal, making it difficult to say if there actually is a significant difference between the two conditions.

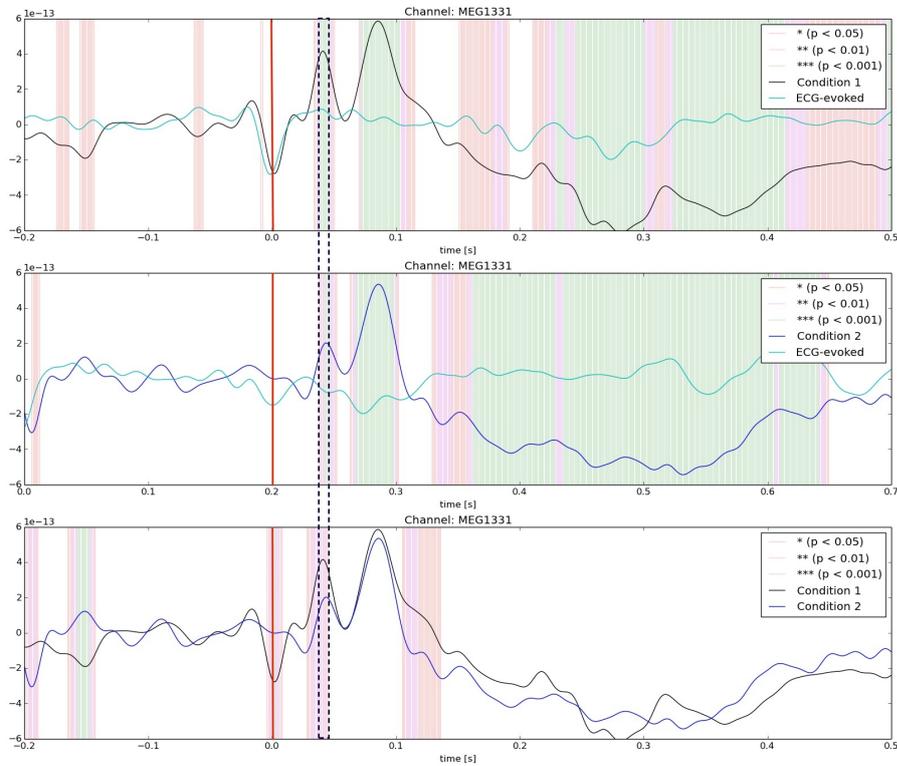


Figure 4.9: The three plots show two-tailed t-test of condition 1 vs. ECG-evoked signal, condition 2 vs. ECG-evoked signal and condition 1 vs. condition 2. All done on data from a single magnetometer located close to SI on subject 2. The red lines indicate the time of stimuli and the black dashed lines mark the time span of interest (where a significant difference can be seen in all three plots).

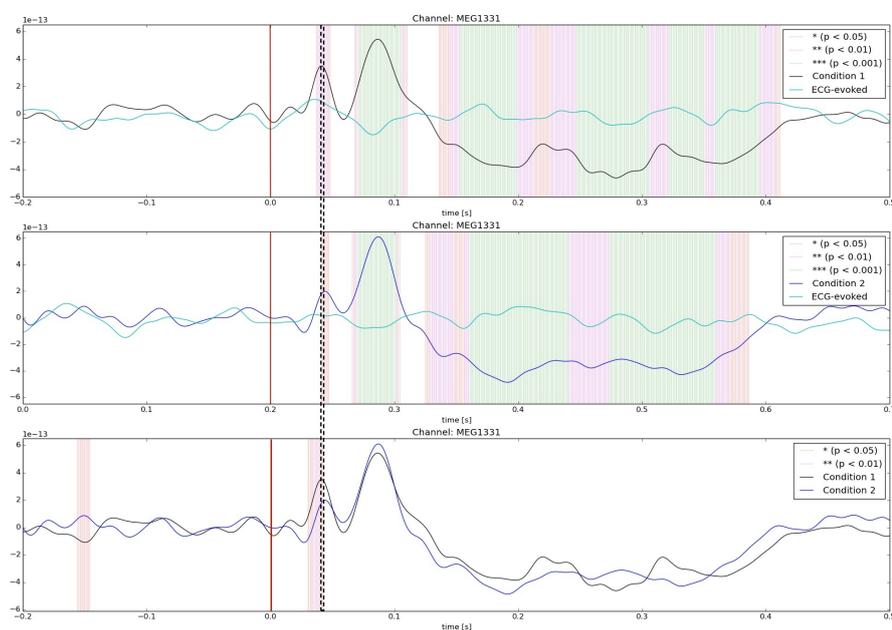


Figure 4.10: The three plots show two-tailed t-test of condition 1 vs. ECG-evoked signal, condition 2 vs. ECG-evoked signal and condition 1 vs. condition 2. All done on ICA processed data from a single magnetometer located close to SI on subject 2. The red lines indicate the time of stimuli and the black dashed lines mark the time span of interest (where a significant difference can be seen in all three plots).

4.1.4 Source Level

Although more effort was put into developing the sensor level part of the analysis pipeline, scripts for source level analysis were also written. The processing steps needed for source level estimates (including BEM, setting up source space, co-registration, forward model, noise covariance matrix and inverse operator) were executed without errors. However, it is difficult to determine whether the co-registration of MEG and MRI coordinate systems was done in the best possible way, since it is a processing step best performed by an experienced user due to the fact that it's mainly based on visual inspection. The first activity peak on source level using dSPM can be seen in Figure 4.11, with the activity peaking in an area corresponds to the S1 region in the right hemisphere as expected from theory as well as from sensor level analysis (compare to Figure 4.4(a)), indicating that the analysis pipeline works.

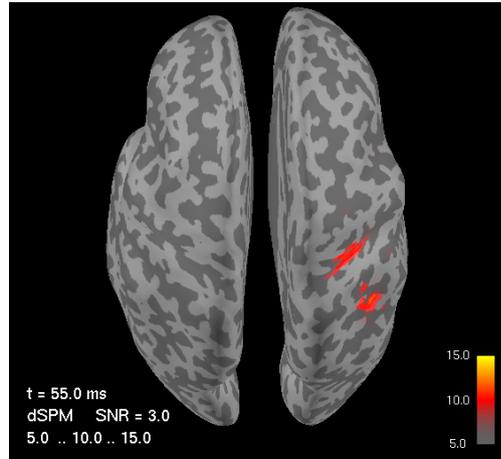


Figure 4.11: The brain of subject 1 (inflated to facilitate viewing of activity in sulci) from MNE GUI *mne_analyze*. An activation around S1 is noticeable 55 ms after stimulus, using dSPM.

4.2 Testing on New Data Sets

Data from measurements on the two subjects previously categorized with microneurography (subject 3 and 4) was used for testing the developed analysis pipeline. It enabled a better understanding of the parts of the pipeline that would benefit from being modified, and it also provided more data, improving the understanding of how the results of different processing steps may vary depending on subject and measurement conditions.

When doing the new measurements, minor adjustments were made to the experimental setup as compared to the two pilot measurements, from which the analysis pipeline was developed. The data channel previously used for recording the stimuli now also recorded each R-wave of the ECG, each stimulus as well as a feedback signal of each stimulus. All these events had an event-specific amplitude, so called event number, on the channel. This meant that the previous processing paradigm using every nonzero value on this channel as a marker of a stimulus event no longer worked. The affected scripts were edited according to the new conditions and worked without errors thereafter, producing the expected plots in MNE Report.

The new use of this data channel can in the future lead to a simplification of the analysis scripts, since much important information is gathered in one channel. The new experimental setup leads to different event numbers for an R-wave occurring without an applied stimulus, an R-wave with stimulus (condition 1) and stimulus without an R-wave (condition 2). The reading of an excel sheet for division of conditions is thereby unnecessary and the data channel can also be used for finding ECG-epochs without the need of manually separating the ones coinciding with stimulus events.

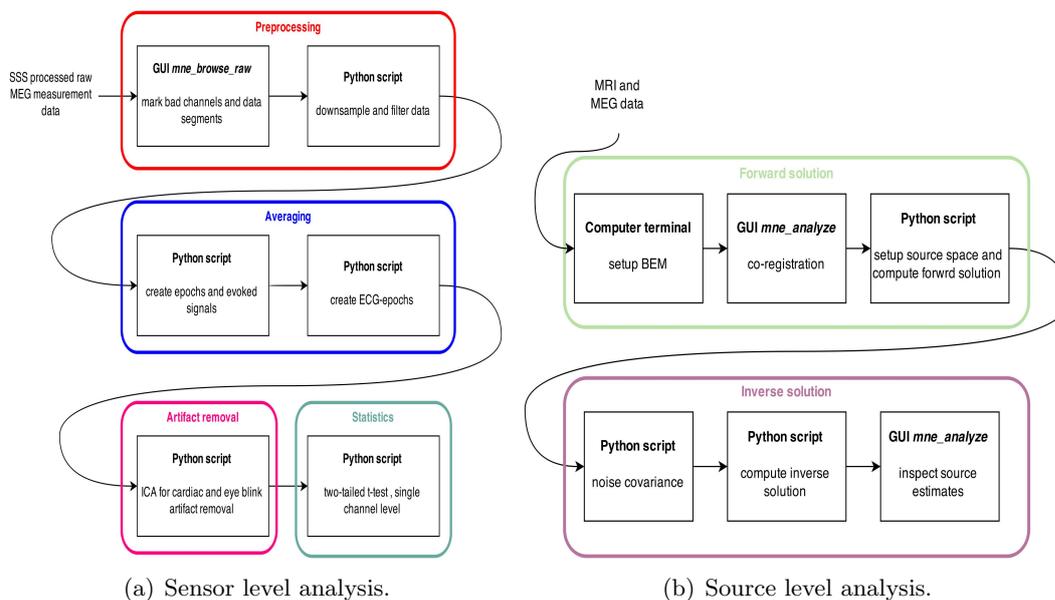


Figure 4.12: Analysis pipeline for MEG measurement data. Each box represents one analysis script or user performed processing step. Further description can be found in Table 4.1 and 4.2.

4.3 Analysis Pipeline

The basic flow chart of the developed analysis pipeline is presented in Figure 4.12, divided into sensor and source level processing. A more detailed description of the processing steps in the analysis pipeline can be found in Table 4.1 and Table 4.2, where it can be seen that the majority of processing steps are performed by running separate python scripts for separate steps, enabling changing of parameters in one processing step without having to run a complete analysis. The python scripts for complete sensor level analysis (except for statistical analysis) together with the produced MNE Report enabling a quick overview of a data set can be found in Appendix A.

The scripts are written in such a way that the only thing the user needs to specify in order to run the analysis is the subject name (provided that the data sets are located in the standard folder named according to convention), but parameters can easily be changed in the analysis scripts according to user preferences.

Table 4.1: Sensor level analysis divided into summarized processing steps including information about what the user should do manually and/or look for.

Sensor Level	
Preprocessing	
<i>mne_browse_raw</i>	USER: Visually inspect the raw data, look for and mark bad channels and data segments, save user defined event list containing bad data segments.
Python	SCRIPT: Downsample to 300 Hz, band-pass filter between 0.6–45 Hz and notch filter 50, 100 and 150 Hz. Plot PSD before and after filtering and save downsampled and filtered data. USER: Inspect the results of filtering on PSD plots. Modify downsampling and filtering frequencies according to preferences.
Averaging	
Python	SCRIPT: Create epochs around each stimulus event with the time span $(-0.25, 0.5)$ s as compared to stimulus. Average to compute three evoked signals: one for each condition and one combined. The conditions are separated by reading an excel sheet containing the stimulus protocol, and any epochs coinciding with bad data segments are discarded. The evoked signals and topographical maps at time 50, 100 and 150 ms after stimulus are plotted. Epochs and evoked signals are saved. USER: inspect plots. Modify time spans for epochs and times for topographical plots according to preferences.
Python	SCRIPT: Create ECG-epochs with the time span $(-0.25, 0.5)$ s as compared to the R-wave. Any epochs coinciding with stimulus events and bad data segments are discarded. ECG-epochs and ECG event list are saved.
Artifact Removal	
Python	SCRIPT: FastICA for cardiac and eye blink artifact removal, with CTPS and Pearson correlation respectively for picking the ICs best corresponding to the artifacts. Component and overlay plots are created. The computed ICs as well as ICA processed stimulus epochs, stimulus evoked and ECG-epochs are saved. USER: Inspect all plots, evaluate if the correct ICs have been removed from data. Change ICA method and choice of ICs to remove from the data according to preferences.
Statistics	
Python	SCRIPT: Compute and plot two-tailed t-tests for statistical comparison of the two conditions. USER: Choose channel to be plotted and specify if ICA processed data should be used.

Table 4.2: Source level analysis divided into summarized processing steps including information about what the user should do manually and/or look for.

Source Level	
Forward solution	
Computer terminal	USER: Setup boundary element model using MRI data and FreeSurfer software.
<i>mne_analyze</i>	USER: Do co-registration of MRI and MEG data coordinate systems and save the .trans-file containing the coordinate transformation.
Python	SCRIPT: Setup source space with spacing <i>oct6</i> corresponding to a distance of approximately 3.1 mm between sources and compute forwards solution. Source space and forward solution are saved.
Inverse solution	
Python	SCRIPT: Compute and save noise covariance from time span $(-0.175, -0.25)$ s as compared to stimulus.
Python	SCRIPT: Compute and save inverse operator with default settings.
<i>mne_analyze</i>	USER: Inspect source estimates.

5

Conclusion

An analysis pipeline for MEG measurement data using the MNE Software was developed. It includes the sensor level processing steps preprocessing, averaging, artifact removal with ICA and statistical analysis as well as the source level analysis steps forward and inverse solution, enabling estimation of the neuronal activity.

The pipeline was developed using two pilot measurement data sets within a medical research project, and after minor adjustments to the analysis scripts due to a small alteration of the experimental setup, the developed pipeline was able to process the data from two new data sets without giving any errors. Expected results in terms of finding a clear, early response to the stimulus in the right hemisphere was produced for all four subjects on sensor level as well as on source level for the two pilot data sets on which source level analysis was done. This indicates that the analysis pipeline works.

5.1 Limitations

There are a few factors limiting the level at which the development of the pipeline could be done related to the specific experimental protocol, the limited amount of data and lack of prior experience.

- The experimental protocol used (giving an arousal stimulus time-locked with the heartbeat) will always produce data containing physiological artifacts related to heartbeats and eye blinks. The artifact problem wasn't expected when starting the analysis protocol development but was considered an important aspect and therefore slightly shifted the focus from source level analysis to sensor level analysis. Although SSS processed data was combined with inbuilt functions in the MNE software package for artifact removal, the artifacts weren't removed completely, making it difficult to determine how much they affect data.
- Having only two data sets when developing the pipeline limited the kind of noise and artifacts seen. For example, only one of the two pilot measurement data sets

used for the development contained EOG measurements, leaving only one data set on which eye blink artifact removal methods could be tested. The total number of data sets on which the pipeline has been tested are four, making it difficult to have sufficient statistics on how the different artifact removal methods affected the data.

- This thesis work was done in collaboration with a PhD student. While neither of us had any prior experience in handling MEG data, we managed to get familiar with many concepts related to MEG and processing of MEG data, and developed an analysis pipeline. Our lack of experience meant that some processes might have been more time demanding than for experienced MEG and MNE users, e.g., processing steps requiring experience like co-registration and visually determine the look of artifacts, components to remove etc.

5.2 Future Aspects

The main thing to do to complete the processing pipeline in its current state is to simplify the scripts according to the new experimental setup. Thereafter, the plan is to present the Arousal Project experimental protocol and setup together with the developed analysis pipeline to MEG experts in order to get more extensive feedback. Other things to consider for future development of the pipeline are listed below.

- Adding frequency analysis to the protocol, which currently only contains time domain analysis. Filtering, downsampling and averaging of data needs to be re-considered if frequency analysis is to be done.
- For further comparison of different conditions the statistical analysis should be further investigated, e.g., instead of looking at one channel at a time, and average over a number of channels showing a specific response can be used.
- To further investigate the measurement data the focus can be shifted towards source level analysis, now that the sensor level pipeline is developed with much detail. This would enable utilizing more of the potential of MEG by investigating the neuronal activity in different brain regions.
- When more data is available (after further measurements have been done), group-level analysis comparing responders to non-responders can be done.
- Other MEG data processing tools will also be investigated, probably starting with FieldTrip, to see if for example artifact removal can be done in a better way.

Bibliography

- [1] R. Wilkins, *Neurosurgical Classics*, American Association of Neurological Surgeons, Thieme (1992) 1.
- [2] A. Debernardi, E. Sala, G. D’Aliberti, G. Talamonti, A. Franchini, M. Collice, *Alcmaeon of Croton, Neurosurgery* 66 (2) (2010) 247–252.
- [3] Human Brain Project.
URL <https://www.humanbrainproject.eu/>
- [4] S. Supek, A. C.J, *Magnetoencephalography From Signals to Dynamic Cortical Network*, Springer 55 (2014) 1.
- [5] A. Gramfort, M. Luessi, E. Larson, D. Engemann, S. Stroheimer, C. Brodbeck, R. Goj, M. Jas, T. Brooks, L. Parkkonen, M. Hämäläinen, *MEG and EEG Data Analysis with MNE-Python*, *Frontiers in Neuroscience* 7 (6) (2013) 1–13.
- [6] M. Hämäläinen, R. Hari, I. Risto, J. Knuutila, O. Lounasmaa, *Magnetoencephalography - Theory, Instrumentation, and Applications to Noninvasive Studies of the Working Human Brain*, *Reviews of Modern Physics* 65 (2) (1993) 413–497.
- [7] R. Hari, R. Salmelin, *Magnetoencephalography: From SQUIDS to neuroscience. NeuroImage 20th Anniversary Special Edition.*, *NeuroImage* 61 (2012) 386–396.
- [8] J. Schneiderman, *Information Content with Low- vs. High- T_C SQUID Arrays in MEG Recordings: The Case for High- T_C SQUID-Based MEG*, *Journal of Neuroscience Methods* 222 (2014) 42–46.
- [9] V. Donadio, R. Liguori, M. Elam, T. Karlsson, M. Giannoccaro, G. Pegenius, F. Giambattistelli, B. Wallin, *Muscle Sympathetic Response to Arousal Predicts Neurovascular Reactivity During Mental Stress*, *The Journal of Physiology* 590 (12) (2012) 2885–2996.
- [10] V. Donadio, T. Karlsson, M. Elam, B. Wallin, *Interindividual Differences in Sympathetic and Effector Responses to Arousal in Humans*, *Journal of Physiology* 544 (1) (2002) 293–302.

BIBLIOGRAPHY

- [11] ©User:Sebastian023 / Wikimedia Commons / CC-BY-SA-3.0.
URL <http://commons.wikimedia.org/wiki/File:LobesCptsLateral.png>
- [12] P. Schimpf, C. Ramon, J. Haueisen, Dipole Models for the EEG and MEG, *IEEE Transactions on Biomedical Engineering* 49 (2) (2002) 409–418.
- [13] Tom Holroyd / Wikimedia Commons / Public Domain.
URL <http://commons.wikimedia.org/wiki/File:LobesCptsLateral.png>
- [14] J. Sarvas, Basic Mathematical and Electromagnetic Concepts of the Biomagnetic Inverse Problem, *Physics in Medicine and Biology* 32 (1) (1987) 11–22.
- [15] C. Lu, J. Huang, S. Chiu, J. Jeng, High-Sensitivity Low-Noise Miniature Fluxgate Magnetometers Usig a Flip Chip Conceptual Design, *Sensors* 14 (8) (2014) 13815–13829.
- [16] J. Clark, M. Neuman, W. Wolson, R. Peura, F. Primiano, M. Siedband, J. Webster, L. Wheeler, *Biomedical Instrumentation: Application and Design*, John Wiley & Sons, Inc. 4th edition (2010) 181.
- [17] D. Dougherty, S. Rauch, J. Rosenbaum, *Essentials of Neuroimaging for Clinical Practise*, American Psychiatric Publishing Inc. 55 (2004) 22.
- [18] V. Lee, *Cardiovascular MRI: Physical Principles to Practical Protocols*, Lippincott Williams & Wilkins 55 (2006) 15.
- [19] ©User:Herbertweidner / Wikimedia Commons.
URL https://commons.wikimedia.org/wiki/File:SQUID_de.png
- [20] M. Schmelz, R. Stolz, V. Zakosarenko, T. Schönau, S. Anders, L. Fritzsche, M. Mück, H.-G. Meyer, Field-stable SQUID Magnetometer with Sub-fT Hz^{-1/2} Resolution Based on Sub-micrometer Cross-type Josephson Tunnel Junctions, *Superconductor Science and Technology* 24 (2011) 1–5.
- [21] Enthought Canopy.
URL <https://www.enthought.com/products/canopy/>
- [22] FreeSurfer.
URL <http://freesurfer.net>
- [23] M. Hämäläinen, *MNE Software User's Guide, Version 2.7.3*, Martians Center for Biomedical Imaging 16th edition.
- [24] J. Dammers, M. Schiek, F. Boers, C. Silex, M. Zvyaginstev, U. Petrzyk, K. Mathiak, Integration of Amplitude and Phase Statistics for Complete Artifact Removal in Independent Components of Neuromagnetic Recordings, *IEEE Transactions on Biomedical Engineering* 55 (10) (2008) 2353–2362.

- [25] S. Taulu, J. Simola, M. Kajola, Applications of the signal space separation method, *IEEE Transactions on Signal Processing* 53 (9) (2005) 3359–3372.
- [26] S. Taulu, J. Simola, Spatiotemporal signal space separation method for rejecting nearby interference in MEG measurements, *Physics in Medicine and Biology* 51 (2006) 1–10.
- [27] M. Uusitalo, R. Ilmoniemi, Signal-Space Projection Method for Separating MEG or EEG into Components, *Medical and Biological Engineering and Computing* 35 (2) (1997) 135–140.
- [28] A. Hyvärinen, E. Oja, Independent Component Analysis: Algorithms and Applications, *Neural Networks* 13 (4-5) (2000) 411–430.
- [29] A. Gramfort, M. Luessi, E. Larson, D. Engemann, S. Stroheimer, C. Brodbeck, L. Parkkonen, M. Hämäläinen, MNE Software for Processing MEG and EEG Data, *NeuroImage* 86 (2014) 446–460.

A

Python Script

Here follows a python script for sensor level analysis of MEG data including downsampling, filtering, averaging and artifact reduction using ICA. The analysis script is meant to provide an overview of the data set with figures from the processing steps presented in a report, which can be seen in Appendix B. The only thing required to change in the scripts when analysing a new data set is the subject name (provided that the files are named according to convention and saved in the standard folder), the other parameters are set to a standard decided in this thesis work, but can easily be changed according to user preferences.

```
# Sensor level analysis script generating a report.
# Including downsampling, filtering, averaging, ECG and EOG
# artifact removal with ICA

import mne
from mne import io
import numpy as np
import xlrtd
import os.path
from mne.preprocessing import ICA, create_eog_epochs
from mne.report import Report

#####
# Set parameters

Subject = 'Subject1'

data_path = "/Users/mtw/Documents/MEG_Data/" + Subject + "/" + Subject
raw_fname = data_path + '_raw_sss.fif'
raw = io.Raw(raw_fname, preload=True)
report = Report()

#####
#Parameters to be Specified
```

```
# Downsampling frequency
DownSamplingfreq = 300

# Data path for excel sheet containing stimulus protocol
data_path_for_xls='/Users/mtw/Documents/MEG_raw_data/delay.xls'

# Band-pass frequencies
lfreq = 0.6
hfreq = 45
# Notch filter frequency
powerlinefreq = 50
# Frequencies for PSD plots
fmin=0
fmax=200

# Specify epoch time (trial length)
tmin = -0.25
tmax = 0.5

# Specify stim channel
stim_channel = 'STI101'
# Specify ECG channel
ecg_channel = 'MISC011'

#####

# Assign 'eog = True' to calculate ICA for EOG (only if EOG channels
# are present in data)
eog = True

# Assign 'apply_ICA = true' to apply ICA to epochs and evoked signals and save
# them as well as the computed ICs
apply_ICA = True

#####

# Keep record of original sampling frequency
ActualFreq = raw.info['sfreq']

# Extract event list before downsampling
events = mne.find_events(raw, stim_channel=stim_channel, verbose=True)

# Downsampling the data for fast processing
if(ActualFreq-DownSamplingfreq>0):
    raw.resample(DownSamplingfreq)

# Plot PSD before filtering
fig1 = raw.plot_psd(area_mode='range')

# Notch + band-pass filtering
raw.notch_filter(np.arange(powerlinefreq, DownSamplingfreq/2, powerlinefreq),
```

APPENDIX A. PYTHON SCRIPT

```
n_jobs=1)
raw.filter(l_freq=l_freq, h_freq=h_freq, method='iir')

# Plot PSD after filtering
fig2 = raw.plot_psd(fmin=fmin, fmax=fmax, area_mode='range')

fig = [fig1, fig2]
report.add_section(fig, captions=['Downsampled data',
                                  'Downsampled, notch and band-pass filtered data'],
                  section='PSD Plots')

#####
# SAVE event list and filtered data
mne.write_events(data_path+'-StimEvents-eve.fif', events)
raw.save(data_path + '_filt-point6-45_raw_sss.fif', overwrite=True)

#####
# Load user defined events (containing bad data segments)

# Make sure the user defined event files exist
if(os.path.exists(data_path+'-raw-eve.fif')):
    events_1 = mne.read_events(data_path+'-raw-eve.fif')
else:
    events_1 = []
if(os.path.exists(data_path+'-raw1-eve.fif')):
    events_2 = mne.read_events(data_path+'-raw1-eve.fif')
else:
    events_2 = []
if(os.path.exists(data_path+'-raw2-eve.fif')):
    events_3 = mne.read_events(data_path+'-raw2-eve.fif')
else:
    events_3 = []
if(os.path.exists(data_path+'-raw3-eve.fif')):
    events_4 = mne.read_events(data_path+'-raw3-eve.fif')
else:
    events_4 = []

# Concatenate user defined events
events_user = np.vstack((events_1,events_2,events_3,events_4))

#####
# Read excel sheet to find the events corresponding to the two conditions
# (stim on R-wave or stim with 200 ms delay) and delete events coinciding with
# bad data segments

workbook = xlrd.open_workbook(data_path_for_xls)
sheet = workbook.sheet_names()
sheetname = sheet.pop()
worksheet = workbook.sheet_by_name(sheetname)

# Loop through cell values, save as array
l = np.zeros(36,np.int8)
for i in range(1,37):
```

```

    s = worksheet.cell_value(i,5)
    l[i-1] = int(s)-1 # -1 to find the correct index in events
l2 = np.zeros(36,np.int8)
for i in range(1,37):
    s = worksheet.cell_value(i,6)
    l2[i-1] = int(s)-1

# Assign event id 1 (for R-wave condition) and 2 (for 200 ms delay condition)
for i in range(0,36):
    events[l[i],2] = 1
    events[l2[i],2] = 2

# Drop any stim events coinciding with bad data segments
flag = events.size/3
s1 = events.size/3 #number of stim events
s2 = events_user.size/3 #number of user defined events

for i in range (0, s2):
    for j in range (0,s1):
        if(abs(events[j][0]-events_user[i][0]) < 0.300*ActualFreq):
            events = np.delete(events,j,axis=0)
            s1 = s1-1
            break

print 'Deleted number of events from Stim channel are', flag-s1

#####

# Resample event lists according to new sampling frequency (downsampled data)
if(ActualFreq-DownSamplingfreq>0):
    events[:,0] = events[:,0]*DownSamplingfreq/ActualFreq
if(events_user.size):
    events_user[:,0] = events_user[:,0]*DownSamplingfreq/ActualFreq

# Define variables needed to create epochs
event_id = dict(ms0=1, ms200=2)
baseline = (-0.1, -0.025)
picks = mne.pick_types(raw.info, meg=True, eeg=False, eog=True,
                        exclude='bads')

reject = dict()

# Create epochs
epoch = mne.Epochs(raw, events, event_id, tmin, tmax, proj=True,
                   picks=picks, baseline=baseline, reject=reject,
                   preload=True,verbose=False)

# Equalize the event counts to maintain the same number of events
# for both conditions
epoch.equalize_event_counts(['ms0','ms200'], method="mintime")

print 'Number of trials for each condition are' , epoch['ms0'].events.shape[0]

```

```

Count=epoch['ms0'].events.shape[0]

# Create evokeds (list containing evoked arrays for three conditions:
# 0ms, 200ms and combined)
evokeds = [epoch[name].average() for name in 'ms0', 'ms200']
evoked = epoch.average()
evokeds.append(evoked)

# Plot evoked butterfly plots
title = dict(grad = 'All conditions - Gradiometers',
             mag = 'All conditions - Magnetometers')
title0 = dict(grad = 'R-wave - Gradiometers',
             mag = 'R-peak - Magnetometers')
title200 = dict(grad = '200 ms delay - Gradiometers',
             mag = '200 ms delay - Magnetometers')

fig = []
fig1 = evokeds[2].plot(titles=title)
fig2 = evokeds[0].plot(titles = title0)
fig3 = evokeds[1].plot(titles = title200)

# Plot evoked topomap plots
fig4 = evokeds[0].plot_topomap(times=np.linspace(0.05, 0.15, 3),ch_type='mag')
fig5 = evokeds[1].plot_topomap(times=np.linspace(0.05, 0.15, 3),ch_type='mag')

# Add to report
fig = [fig1,fig2,fig3,fig4,fig5]
combined = 'Combined Conditions'
cond1 = 'Condition 1 - R-wave'
cond2 = 'Condition 2 - 200 ms delay'
report.add_section(fig, captions=[combined,cond1,cond2,cond1,cond2],
                  section='Evoked')

#####
# SAVE resampled events, epochs and evokeds

if(events_user.size):
    mne.write_events(data_path+'-'+str(DownSamplingfreq)+'-User-eve.fif',
                    events_user)
mne.write_events(data_path+'-'+str(DownSamplingfreq)+'-Stim-eve.fif', events)
epoch.save(data_path+'-epo.fif')
mne.write_evokeds(data_path+'-ave.fif', evokeds)

#####
# Create ECG-epochs not coinciding with stim events or bad data segments

# Combine event lists (if user defined events exist)
if(events_user.any()):
    events_com = np.vstack((events,events_user))
else:
    events_com = events

# Sort the combined event list according to time

```

```

events_com = events_com[events_com[:,0].argsort()]

# Pick channel types for ECG epochs
picks = mne.pick_types(raw.info, meg=True, eeg=False, eog=False,
                       stim=False, exclude='bads')

# Find ECG events
ecg_epochs_events,_,_ = mne.preprocessing.find_ecg_events(raw,
                                                          event_id=999, ch_name=ecg_channel)

# Drop ECG events coinciding with stim events and bad data segments
j = 0;
drops = 0
print 'Original number of stim events and bad data segments:', events_com.shape[0]
for i in range (0, events_com.shape[0]):
    while((ecg_epochs_events[j][0] < events_com[i][0]) and (j < ecg_epochs_events.shape[0])):
        j = j+1
    if(j+1 >= ecg_epochs_events.shape[0]):
        break
    if(abs(ecg_epochs_events[j-1][0] - events_com[i][0]) <= 0.250*DownSamplingfreq):
        ecg_epochs_events = np.delete(ecg_epochs_events, j-1, axis=0)
        drops = drops+1
    if(abs(ecg_epochs_events[j][0] - events_com[i][0]) <= 0.250*DownSamplingfreq):
        ecg_epochs_events = np.delete(ecg_epochs_events, j, axis=0)
        drops = drops+1
    if(abs(ecg_epochs_events[j+1][0] - events_com[i][0]) <= 0.250*DownSamplingfreq):
        ecg_epochs_events = np.delete(ecg_epochs_events, j+1, axis=0)
        drops = drops+1
print 'Number of dropped ECG-events:', drops

# Define variables needed to create epochs
ecg_events = ecg_epochs_events.copy()
event_id = dict(B=999)

# Create ECG epochs
ecg_epochs = mne.Epochs(raw, ecg_events, event_id, tmin, tmax, proj=True,
                       picks=picks, baseline=baseline, reject=reject,
                       preload=True, verbose=False)

# Create ECG-evoked
ecg_evoked = ecg_epochs.average()

#####
# SAVE ECG-events and ECG-epochs

mne.write_events(data_path+'-ECG-eve.fif', ecg_epochs_events)
ecg_epochs.save(data_path+'-ECG-epo.fif')

#####
# ICA for cardiac artifact

ica = ICA(n_components=0.95, method='fastica')
ica.fit(ecg_epochs)

```

APPENDIX A. PYTHON SCRIPT

```
# Detection of components related to heartbeat via cross trial phase statistics
ecg_inds, scores = ica.find_bads_ecg(ecg_epochs)

# Plot scores
title = 'Sources related to %s artifacts (red)'
fig1 = ica.plot_scores(scores, exclude=ecg_inds, title=title % 'ecg')

# Pick the five largest scores and plot them
show_picks = np.abs(scores).argsort()[::-1][:5]

# Plot sources
fig2 = ica.plot_sources(epoch, show_picks, exclude=ecg_inds, title=title % 'ecg')

# Topomap plots of chosen components
fig3 = ica.plot_components(ecg_inds, title=title % 'ecg', colorbar=True)

# Exclude the components related to heartbeat
n_max_ecg = 5 # Maximum nbr of components to reject
ecg_inds = ecg_inds[:n_max_ecg]
ica.exclude = ecg_inds

#####
# Assess component selection by visual inspection of plots

# Plot ECG sources + selection
fig4 = ica.plot_sources(ecg_evoked, exclude=ecg_inds)

# Plot ECG-evoked overlay
fig5 = ica.plot_overlay(ecg_evoked, exclude=ecg_inds)

# Plot evoked overlay for R-wave, 200 ms delay and combined events
fig6=ica.plot_overlay(evokeds[2])
fig7=ica.plot_overlay(evokeds[0])
fig8=ica.plot_overlay(evokeds[1])

# Add to report
fig=[]
fig=[fig1,fig2,fig3,fig4,fig5,fig6,fig7,fig8]
report.add_section(fig, captions=['Scores of ICs related to ECG',
    'Time Series plots of ICs (ECG)', 'TopoMap of ICs (ECG)',
    'Time-locked ECG sources', 'ECG overlay',
    'Combined conditions overlay (ECG)', 'Condition 1 Overlay (ECG)',
    'Condition 2 overlay (ECG)'], section='ICA - ECG')

#####
# ICA for eye blink artifact

if (eog):
    n_max_eog = 2
    eog_inds, scores = ica.find_bads_eog(raw)
```

```

scores = scores[1]
fig1 = ica.plot_scores(scores, exclude=eog_inds, title=title % 'eog')
show_picks = np.abs(scores).argsort()[::-1][:5]
fig2 = ica.plot_sources(raw, show_picks, exclude=eog_inds, title=title % 'eog')
fig3 = ica.plot_components(eog_inds, title=title % 'eog', colorbar=True)
eog_inds = eog_inds[:n_max_eog]
ica.exclude += eog_inds

picks = mne.pick_types(raw.info, meg=True, eeg=False, eog=False,
                       stim=False, exclude='bads')
eog_evoked = create_eog_epochs(raw, tmin=-.5, tmax=.5, picks=picks).average()
fig4 = ica.plot_sources(eog_evoked, exclude=eog_inds)
fig5 = ica.plot_overlay(eog_evoked, exclude=eog_inds)
fig6 = ica.plot_overlay(evokeds[2])
fig7 = ica.plot_overlay(evokeds[0])
fig8 = ica.plot_overlay(evokeds[1])
fig=[]
fig=[fig1,fig2,fig3,fig4,fig5,fig6,fig7,fig8]
report.add_section(fig, captions=['Scores of ICs related to EOG',
                                  'Time Series plots of ICs (EOG)', 'TopoMap of ICs (EOG)',
                                  'Time-locked EOG sources', 'EOG overlay',
                                  'Combined conditions overlay (ECG+EOG)',
                                  'Condition 1 Overlay (ECG+EOG)',
                                  'Condition 2 overlay (ECG+EOG)'], section='ICA - EOG')

#####
# Apply ICA to evokeds and epochs

if (apply_ICA):
    # Apply ICA to evoked
    ica.apply(evokeds[0])
    ica.apply(evokeds[1])
    ica.apply(evokeds[2])

    # Apply ICA to epochs
    ica.apply(epoch)

#####
# SAVE ICA, ICA-epochs and ICA-evoked
    ica.save(data_path+'-ica.fif')
    epoch.save(data_path+'-ICA-epo.fif')
    mne.write_evokeds(data_path+'-ICA-ave.fif', evokeds)

report.save(data_path+'report.html', overwrite=True)

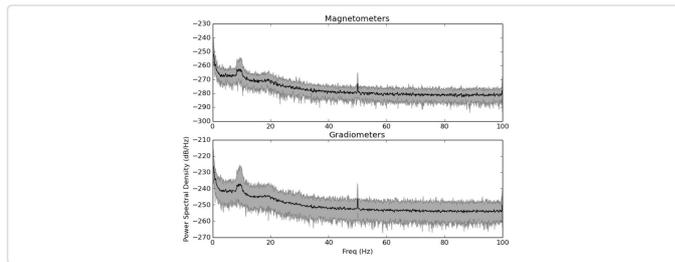
```

B

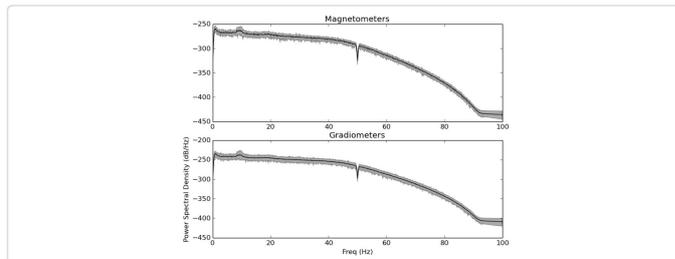
Generated Report

MNE Report created by script in Appendix A for subject 1.

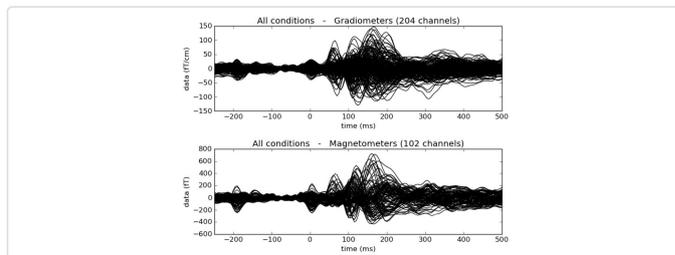
Downsampled data



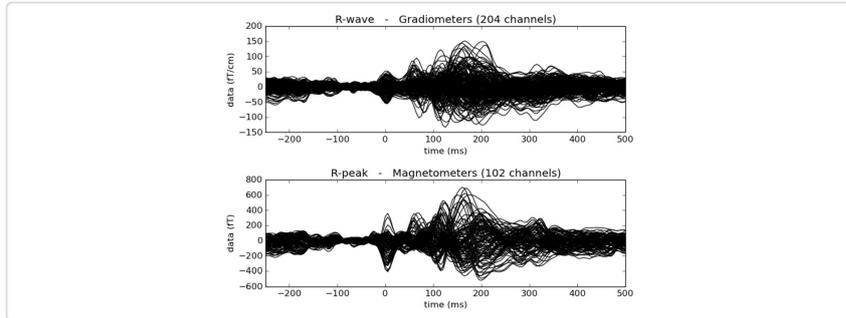
Downsampled, notch and band-pass filtered data



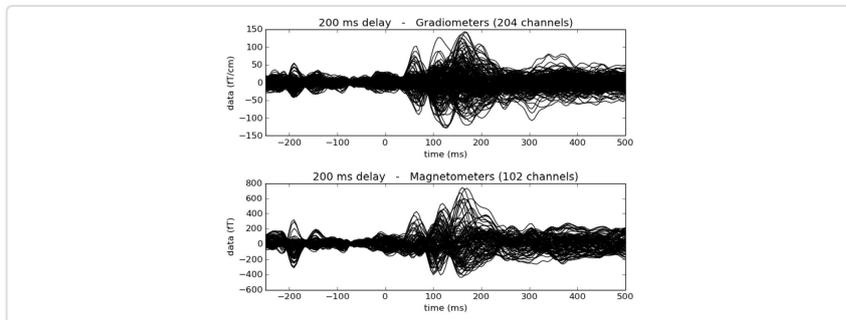
Combined Conditions



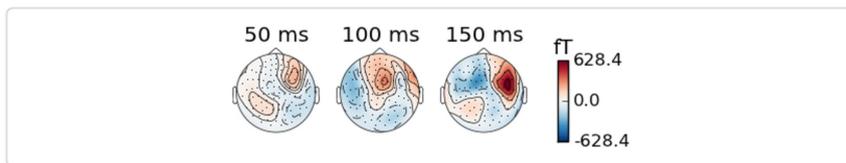
Condition 1 - R-wave



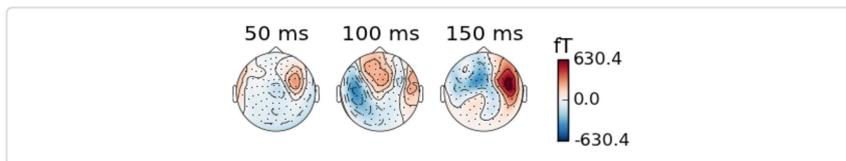
Condition 2 - 200 ms delay



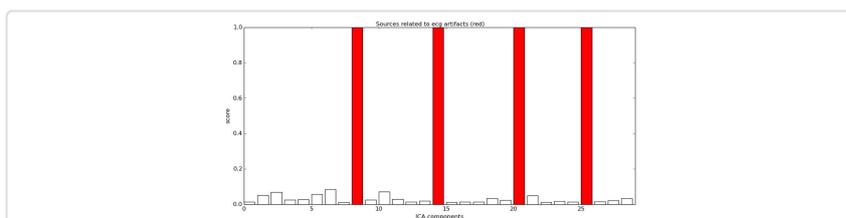
Condition 1 - R-wave



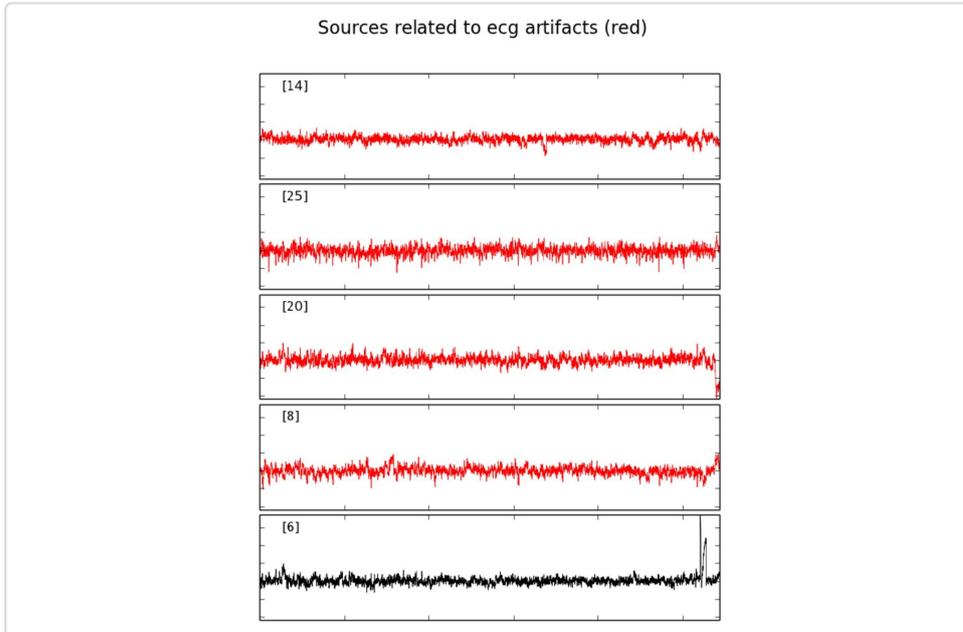
Condition 2 - 200 ms delay



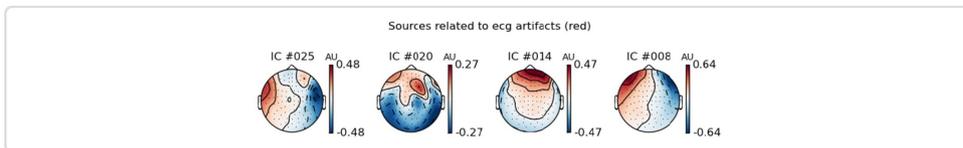
Scores of ICs related to ECG



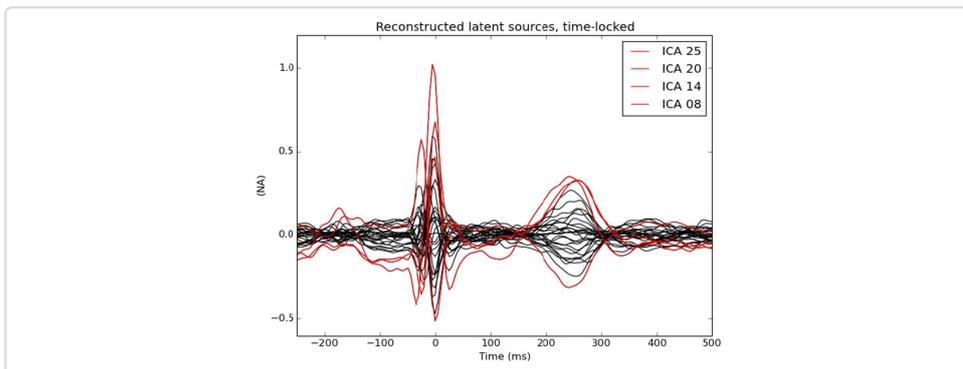
Time Series plots of ICs (ECG)



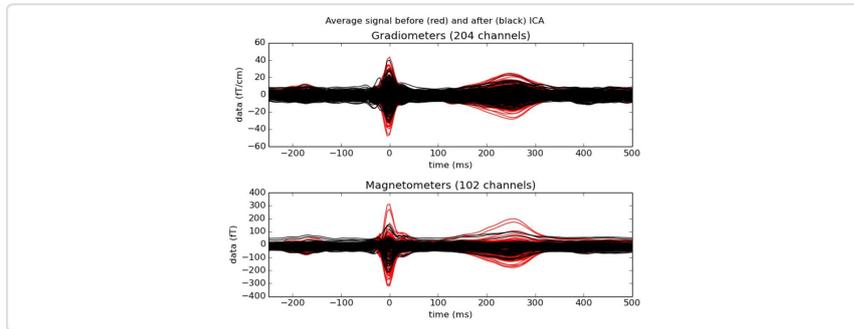
TopoMap of ICs (ECG)



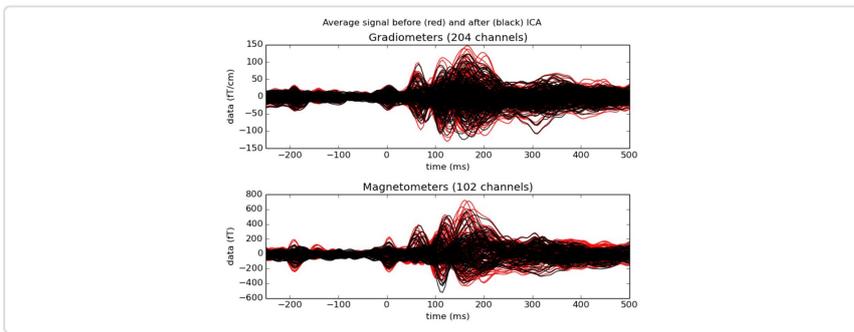
Time-locked ECG sources



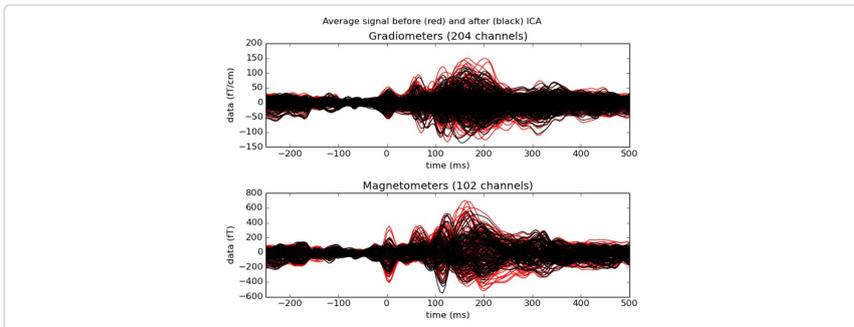
ECG overlay



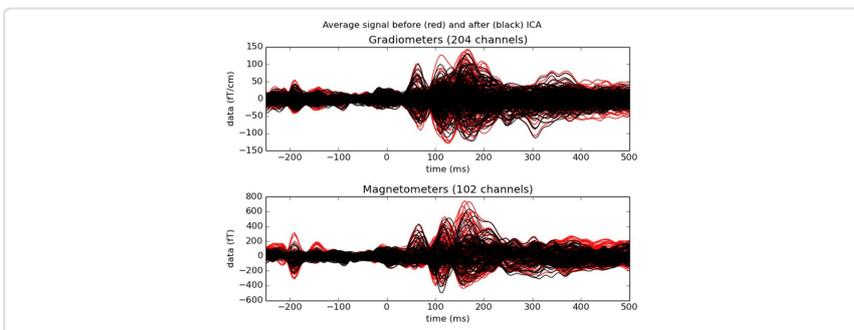
Combined conditions overlay (ECG)



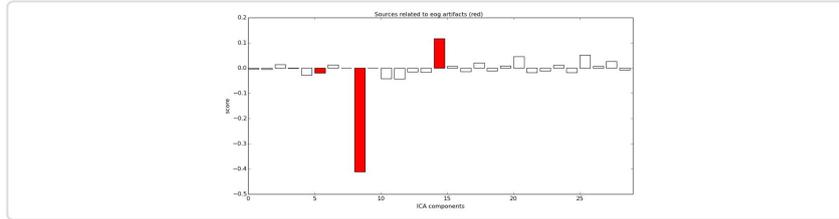
Condition 1 overlay (ECG)



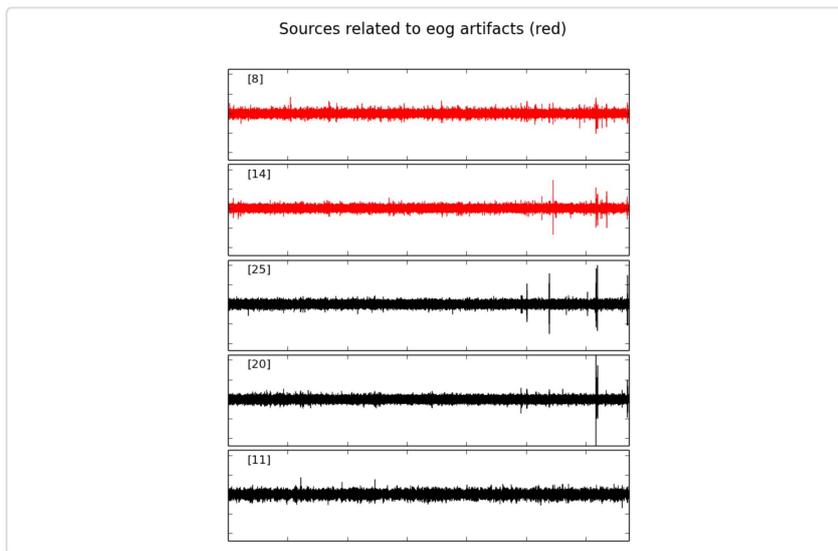
Condition 2 overlay (ECG)



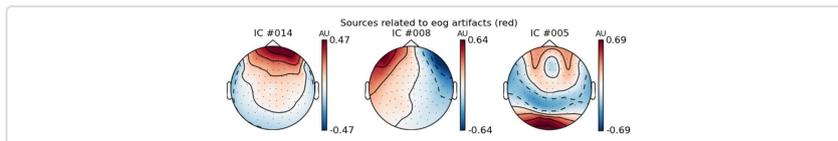
Scores of ICs related to EOG



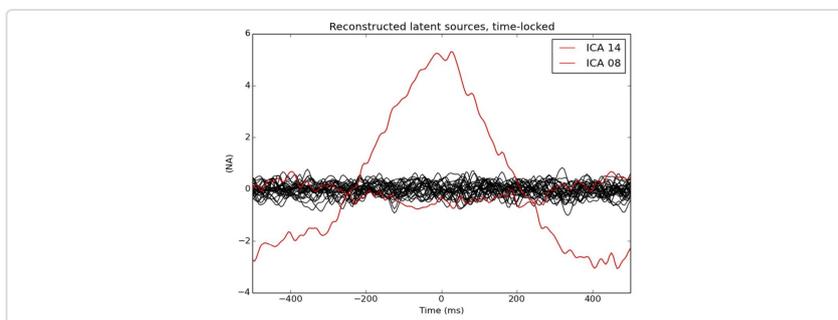
Time Series plots of ICs (EOG)



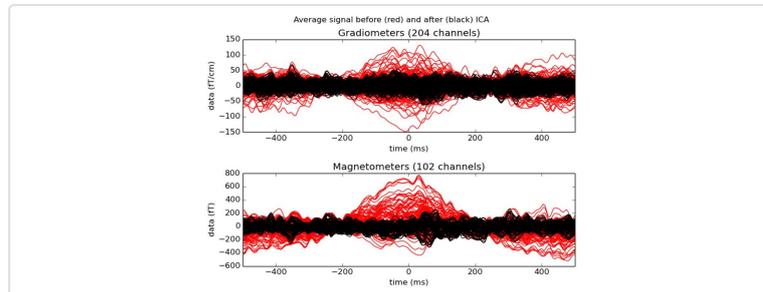
TopoMap of ICs (EOG)



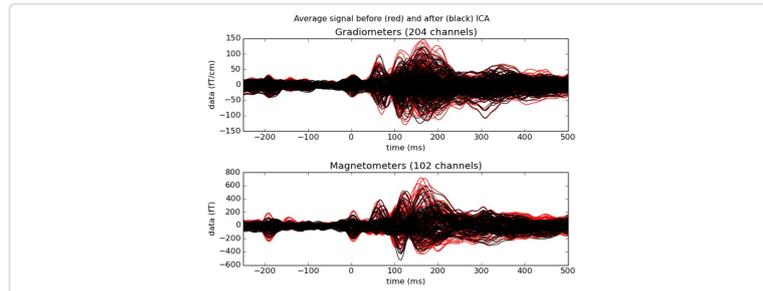
Time-locked EOG sources



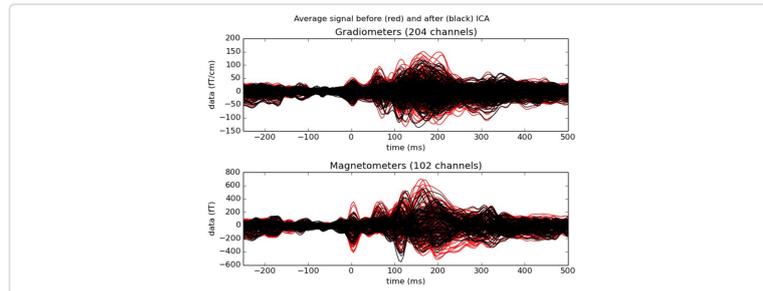
EOG overlay



Combined conditions overlay (ECG+EOG)



Condition 1 overlay (ECG+EOG)



Condition 2 overlay (ECG+EOG)

