UNIVERSIDAD CARLOS III DE MADRID
Departamento de Teoría de la Señal y Comunicaciones

DOCTORAL THESIS

FUNCTIONAL BRAIN IMAGING ON MOBILE DEVICES BY SOLVING THE EEG INVERSE PROBLEM: A STRUCTURED SPARSITY APPROACH

Author: JAIR MONTOYA MARTÍNEZ
Supervised by: ANTONIO ARTÉS RODRÍGUEZ
December 2014

Tesis Doctoral: FUNCTIONAL BRAIN IMAGING ON MOBILE DEVICES BY SOLVING THE EEG INVERSE PROBLEM: A STRUCTURED SPARSITY APPROACH

Autor: Jair Montoya Martínez

Director: D. Antonio Artés Rodríguez

Fecha:

Tribunal

Presidente:

Vocal:

Secretario:

Acknowledgments

First of all, I would like to thank God who has been giving me everything to accomplish this thesis: health, patience, strength, guidance, perseverance and support. I would like to thank my parents, my sister, my brother and my beautiful nieces; thank you so much for all your love, your support, for believing in me and helping to make my dreams come true: this work is for all of you. I want to thank Alexandra, for her love, her patience, her support and for encouraging me to follow my dreams. I wish to express my deepest gratitude to my advisor Professor Antonio Artés Rodríguez for introducing me to the amazing world of machine learning and also for his continuous support and guidance. I would also like to give a special thanks to Professor Massimiliano Pontil, for always being available and sharing his expertise on convex optimization, structured sparsity and matrix factorization, and to Professor Lars Kai Hansen, for introducing me to the exciting world of neuroscience, especially to the EEG inverse problem, and also for giving me the opportunity to visit his lab, Cognitive Systems, at the Technical University of Denmark. During this visit I enjoyed learning how to program Android mobile devices to solve the EEG inverse problem (special thanks to Carsten Stahlhut and Arkadiusz Stopczynski for helping me with my doubts on EEG signal processing and Android programming). I would also like to thank Duver for his great friendship, to Jose Luis, Sandra and Magnus, for making me feel like part of their little family, to Camilo, for sharing with me his passion for maths, to Diego and Raúl, for their spiritual advices, to Juanjo and Alejandro, for our discussions on scientific programming, open source software and swimming techniques, to Marcelino, for always being available to explain/discuss about machine learning and signal processing algorithms. Last but not least, I want to thank my undergrade teacher, Javier Escobar Montoya, may he rest in peace, who made me enjoy and love mathematics.

Abstract

In this thesis we address the development of a mobile brain scanner, which is based on a wireless EEG neuroheadset, in charge of acquiring and transmitting the electrical potential measured on the scalp, and one mobile device (smartphone or tablet), in charge of receiving and processing these data to produce the cortical activation maps, which show, using a 3D brain model, the brain areas that are currently active. To generate the cortical activation maps, the mobile brain scanner needs to solve an electromagnetic inverse problem called the EEG inverse problem. The low spatial resolution of the EEG caused by the low conductivity of the skull plus the small number of EEG sensors available to capture the electrical activity produced by thousands of brain current sources, imply that the EEG inverse problem is underdetermined, ill-posed, and has infinite solutions. To make this problem tractable, in this thesis we assume that the number of active sources is small, that is, we assume that the set of active sources is a sparse set. Additionally, we also assume a linear relationship between the elements of this set. If we represent the set of brain current sources as a matrix (called the sources matrix), where the rows denote how the electrical activity of the sources vary over time, then the former assumptions lead to estimate a sources matrix which is structured sparse and low rank. To solve this problem, in this thesis we propose a method based on the factorization of the sources matrix as a product of two matrices: the first one encodes the spatial dynamics of the sources (how they change their spatial activation patterns), whereas the second one encodes their corresponding temporal dynamics (how they change their electrical activity over time). This method combines the ideas of the Group Lasso (structured sparsity) and Trace Norm (low rank) into one unified framework. We also develop and analyze the convergence of an alternating minimization algorithm to solve the resulting nonsmooth-nonconvex regularization problem. Finally, in order to implement a working prototype of the mobile brain scanner, we bring our method to a real life scenario: online solving of the EEG inverse problem on a mobile device, which is continuously supplied with EEG data coming from the wireless EEG neuroheadset.

Resumen

En esta tesis se aborda el desarrollo de un escáner móvil cerebral, el cual está basado en un casco inalámbrico que captura y transmite señales electroencefalográficas (EEG) a un dispositivo móvil (teléfono inteligente o tableta). Este las recibe y procesa con el fin de generar mapas de activación cerebral, los cuales muestran qué áreas de la corteza cerebral están actualmente activas. Estos mapas son visualizados en la pantalla del dispositivo móvil usando un modelo en 3D del cerebro. Para generar estos mapas de activación, el escáner móvil cerebral debe resolver el problema inverso del EEG. La baja resolución espacial del EEG, causada por la baja conductividad eléctrica del cráneo, añadida al bajo número de sensores EEG disponibles para capturar la actividad eléctrica generada por miles de fuentes cerebrales, hacen que el problema inverso del EEG sea mal condicionado e indeterminado, admitiendo un número infinito de soluciones. Para disminuir la dificultad de este problema, en esta tesis se asume que el número de fuentes eléctricas cerebrales, activadas por un determinado estímulo, es bajo; es decir, se asume que el conjunto de fuentes activas es un conjunto disperso. Adicionalmente, también se asume la existencia de una relación lineal entre los elementos de dicho conjunto. Si se representa el conjunto de las fuentes eléctricas cerebrales usando una matriz, llamada de aquí en adelante matriz de fuentes, las anteriores hipótesis conducen a estimar una matriz de fuentes que sea dispersa, estructurada y de bajo rango. Para resolver este problema, en esta tesis se propone un método basado en la factorización de la matriz de fuentes como el producto de dos matrices: la primera codifica la dinámica espacial de las fuentes (cómo cambian sus patrones de activación), mientras que la segunda incorpora la dinámica temporal de las fuentes (cómo cambian su actividad eléctrica en el tiempo). Este método combina las ideas de dos regularizadores: el regularizador de Grupo Lasso (dispersidad estructurada) y el regularizador de norma nuclear (bajo rango). Para resolver el problema de estimación resultante, el cual es no convexo y no diferenciable, en esta tesis se desarrolla, y también se analiza la convergencia, de un algoritmo de minimización por etapas. Finalmente, para llevar a cabo la implementación de un prototipo funcional del escáner móvil cerebral, se ha trasladado el método propuesto a un escenario de la vida real: solución, en línea, del problema inverso del EEG en un dispositivo móvil, al cual le llegan continuamente datos EEG provenientes del casco inalámbrico de captura de datos.

Contents

Chapter 1

Introduction and Motivation

The brain is the most intricate and complex organ known. With approximately 10^{11} neurons and 10^{15} synaptic connections, it performs computations that exceed the capability of any computer. In order to understand how it works, brain imaging methods combine cognitive science, physics, signal processing and computer science to develop techniques that allow us to see the inner workings of the living brain. These techniques can be categorized as indirect or direct measures of the brain activity. Indirect techniques measure changes in brain properties that are related to neural activity, such as the increased consumption of oxygen in active areas of the brain, whereas direct techniques measure electrical and magnetic fields changes produced by neural action potentials and synaptic activity. The most commonly used indirect techniques are functional Magnetic Resonance Imaging (fMRI) and Positron Emission Tomography (PET). fMRI is a noninvasive technique that measures the brain activity by detecting the changes in the blood oxygenation and flow that occur in response to neural activity. It has high spatial resolution, which allows to obtain detailed activation maps showing which brain regions are involved in a particular cognitive task. However, it has low temporal resolution limited by the slow blood flow response it depends on, which disables it to study the dynamics of the mental activity on the millisecond timescale on which neurons operate. Besides, fMRI requires that the person being scanned stays still inside the MRI scanner, which limits the variety of experiments that can be done (e.g., study the brain under naturalistic conditions). Like fMRI, PET gives access to the neural activity indirectly via the measurement of a metabolic process (the glucose consumption of the different brain areas). It has similar temporal resolution to fMRI, but its spatial resolution is lower. PET is an

invasive technique: it requires the injection of a radiactive tracer, which limits the number of measurements that can be made on the patient.

On the other hand, the most commonly used direct techniques are Magnetoencephalography (MEG) and Electroencephalography (EEG). Both give direct access to neural activity, MEG via the measurement of the magnetic field outside the head, whereas EEG via the measurement of the electrical potential on the scalp. Thanks to the rapid response of the electric and magnetic fields to changes in the neural activity, both have high temporal resolution. EEG has low spatial resolution, mainly because the skull, given its low conductivity, disperses the electrical potential on the scalp. MEG has also low spatial resolution, but it is higher than EEG, because the magnetic field measured outside the head is less sensitive to the conductivity properties of the skull.

Unlike fMRI, PET and MEG, EEG does not require very expensive equipment. In fact, nowadays is possible to find low-cost wireless EEG equipment. In this thesis, we propose to use a low-cost wireless EEG neuroheadset, jointly with a mobile device, to develop a mobile brain scanner that can be used to study the brain on a wide variety of scenarios, in which the person being scanned can move and act under naturalistic conditions. Thanks to the high temporal resolution provided by EEG, this mobile brain scanner can be used to scan the brain on the millisecond timescale on which neurons operate. In this system, the wireless EEG neuroheadset is in charge of acquiring and transmitting the electrical potentials measured on the scalp, and the mobile device (smartphone or tablet), is in charge of receiving and processing these data to produce the cortical activation maps, which show the brain areas that are currently active. To generate the cortical activation maps, the mobile brain scanner needs to solve an electromagnetic inverse problem called the EEG inverse problem. The low spatial resolution of the EEG caused by the low conductivity of the skull plus the small number of EEG sensors available to capture the electrical activity produced by thousands of brain current sources, imply that the EEG inverse problem is underdetermined, ill-posed and has infinite solutions: the electrical potentials measured on the scalp may be generated by an infinite number of different brain current sources distributions. To solve this problem, in this thesis we propose a new method based on regularization theory, which involves the replacement of the original problem with a nearby well-posed problem whose solution approximates the required solution. Solutions developed by this theory are stated in terms of a regularization function, which helps us to select, among the infinite solutions, the one that best fulfills some prescribed assumptions. To reduce the complexity

of this problem, in this thesis we assume that the number of active sources is small, that is, we assume that the set of active sources is a sparse set. Additionally, we also assume a linear relationship between the elements of this set. If we represent the set of brain current sources as a matrix (called the sources matrix), where the rows denote how the electrical activity of the sources vary over time, then the former assumptions lead to estimate a sources matrix which is structured sparse and low rank. To solve this problem, in this thesis we propose a method based on the factorization of the sources matrix as a product of two matrices: the first one encodes the spatial dynamics of the sources (how they change their spatial activation patterns), whereas the second one encodes their corresponding temporal dynamics (how they change their electrical activity over time). We show that the proposed method can be solved efficiently, so that it can be used in the mobile brain scanner (more specifically, in the mobile device), to solve, online, the EEG inverse problem to generate, as fast as possible, the cortical activation maps showing the brain areas currently active.

1.1 Thesis outline

Each one of the concepts and ideas involved in the development of the proposed mobile brain scanner, as well as the ones related with the new method to solve the EEG inverse problem, will be explained in the following chapters:

- **Chapter 2, Anatomical and physiological basis of EEG**, gives an introduction to brain anatomy, neural activity and electroencephalography (EEG). These basic concepts are necessary to understand the behavior and the main features of the brain current sources that we want to induce as solutions of the EEG inverse problem.

- **Chapter 3, EEG forward and inverse modelling**, presents the EEG forward problem, as well as the EEG inverse problem and how they are related through the lead-field matrix. Additionally, it also presents some basic approaches to solve the EEG inverse problem, which is the main problem that we want to solve in this thesis.

- **Chapter 4, Matrix factorization approach to solve the EEG inverse problem**, explains our proposal to solve the EEG inverse problem, which is a new method that takes into account the structured

sparsity and the low rank of the sources matrix. The method is based on the factorization of the sources matrix as a product of a sparse coding matrix and a dense latent source matrix. The structured sparse-low-rank structure is enforced by minimizing a regularized functional that includes the ℓ_{21}-norm of the coding matrix and the squared Frobenius norm of the latent source matrix. We develop an alternating optimization algorithm to solve the resulting nonsmooth-nonconvex minimization problem. We analyze the convergence of the optimization procedure, and we compare, under different synthetic scenarios, the performance of our method with respect to the Group Lasso and Trace Norm regularizers when they are applied directly to the target matrix.

- **Chapter 5, Mobile brain scanners**, gives an introduction to the mobile brain scanning paradigm. It presents the main building blocks of a mobile brain scanner, as well as the main features of the hardware and software components used in this thesis to build a working prototype of such system.

- **Chapter 6, Development of a mobile brain scanner using the matrix factorization approach**, explains how to adapt the batch matrix factorization approach, described in Chapter 4, into an online approach, such that it can be used in a mobile brain scanner to estimate the active sources as fast as possible. It also presents an experimental validation of the proposed approach using the "right index finger tapping experiment", as well as some pictures of a mobile brain scanner prototype running the proposed algorithm on a Samsung Galaxy Note I smartphone and a Nexus 7 tablet.

- **Chapter 7, Conclusion and future work**, presents the conclusion of this thesis and proposes a selection of future research directions.

Chapter 2

Anatomical and physiological basis of EEG

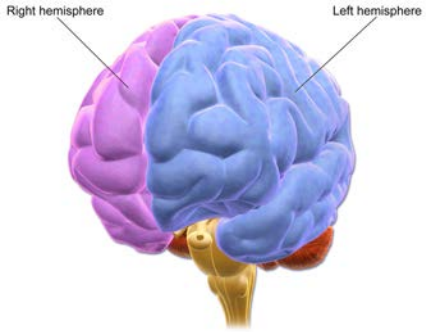
In this chapter fundamental theory related with biological aspects of the physiology and functioning of the human brain will be introduced, in order to understand how the EEG signal is generated, as well as how it can be used to study the brain activity.

2.1 General structure of the human brain

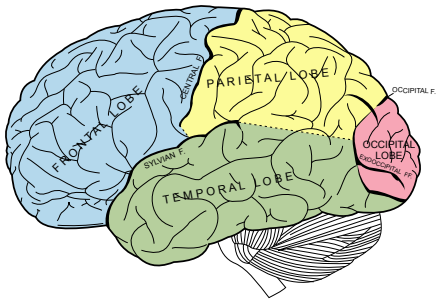
The brain consists of two hemispheres, the right and the left hemispheres (see Figure 2.1(a)), which are separated by the longitudinal fissure. These, in turn, are divided into four lobes (frontal, parietal, temporal and occipital) by two deep grooves: the Rolandic fissure (central fissure), which runs down the side of both hemispheres, and the Sylvian fissure (lateral fissure), which is almost horizontal (see Figure 2.1(b)). Fissures are also commonly called sulci.

The counterpart of the cortical fissures are the gyri. As we can see in Figure 2.2(a), gyri are the convolutions of the surface of the cerebral hemispheres caused by the infolding of the cortex, separated by the fissures. Some of the gyri contain brain regions with known cognitive functions, for instance, Figure 2.2(b) shows the precentral and postcentral gyri, which contains the primary motor cortex (M1) and the primary somatosensory cortex (S1), respectively.

The outermost layer of the brain is known as the cerebral cortex. It is about 2-4 mm thick [Kandel et al., 2000]. It is also referred to as the gray matter as it consists of cell bodies and capillaries and contrasts with

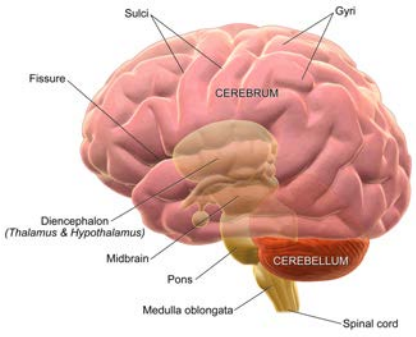


(a) Cerebral hemispheres

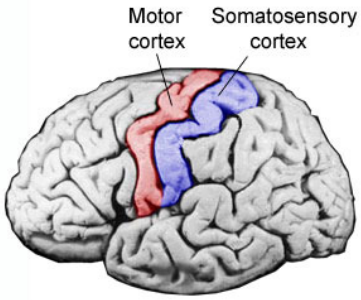


(b) The different lobes of the cerebral cortex

Figure 2.1: Cerebral hemispheres and brain lobes (Source wikipedia.org).



(a) General brain structures including sulci and gyri (Source wikipedia.org)



(b) Motor and Somatosensory cortex (Source http://hopes.stanford.edu/)

Figure 2.2: Cortical fissures and gyri.

the underlying white matter, which consists mainly of the white myelinated sheaths of neural axons, see Figure 2.3.

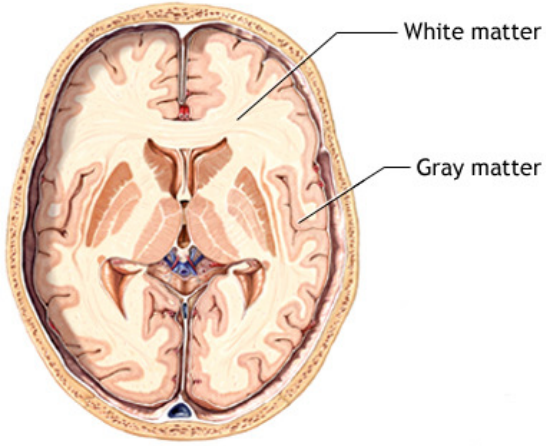


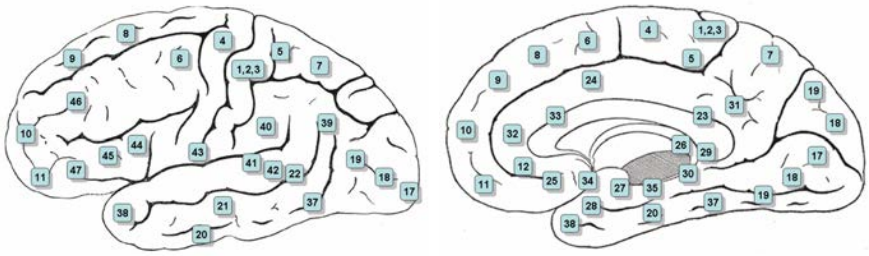
Figure 2.3: Gray and white matter of the brain (Adapted from http://www.nlm.nih.gov/).

The difference on the structure and organization of the cerebral cortex cells led to the neuroanatomist Korbinian Brodmann to divide the cerebral cortex into 52 distintic regions (see Figure 2.4), which are called Brodmann areas [Brodmann, 1909]. Many of the Brodmann areas have been correlated to diverse cortical functions. For example, Brodmann areas 1, 2 and 3 are the primary somatosensory cortex, area 4 is the primary motor cortex, area 17 is the primary visual cortex and areas 41 and 42 correspond closely to primary auditory cortex [Bear et al., 2007].

2.2 Neurons: basic functional units of the brain

From the microscopic point of view, the brain is composed mainly by two types of cell: neurons and glial cells. Neurons act as the information processing units of the brain, while the glial cells ensure the physical structure of the brain, the proper concentration of ions and the transport of the nutrients between the blood vessels and brain tissue [Phillips, 2000].

As we can see in Figure 2.5, the neuron is composed of a soma or cell body, containing the nucleus, a dendritic tree which receives stimuli from other neurons, and the axon, which is the transmitter in charge of carrying the impulse from the soma to the other cells.



(a) Lateral schematic view (b) Medial schematic view

Figure 2.4: Brodmann areas (Source wikipedia.org).

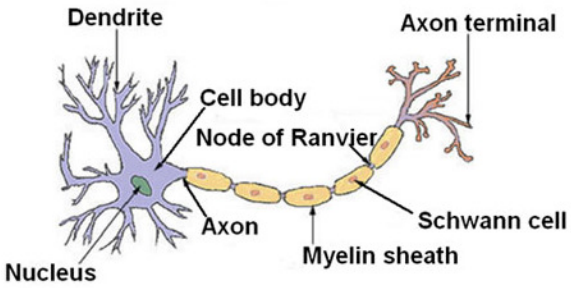


Figure 2.5: Diagram of a neuron (Adapted from wikimedia.org).

The interconnection between neurons, called the synapse, behaves as a simple switch but also has a special role in information processing. The function of the synapse is to transfer electric activity (information) from one cell to another. The transfer can be from nerve to nerve (neuro-neuro), or nerve to muscle (neuro-myo). The region between the pre and postsynaptic membrane is very narrow, only 30-50 nm. It is called the synaptic cleft (or synaptic gap). There is not direct electric communication between the presynaptic and postsynaptic cell; instead, a chemical mediator is utilized [Malmivuo & Plonsey, 1995]. The sequence of events is as follows:

1. An action pulse reaches the terminal endings of the presynaptic cell.

2. A neurotransmitter is released, which diffuses across the synaptic gap to bind to receptors in specialized membranes of the postsynaptic cell.

3. The transmitter acts to open channels of one or several ion species

(Na+ and K+), resulting in a change in the transmembrane potential. This is called a postsynaptic potential (PSP).

4. The resting state potential of the neuron is approximately -70 mV. When a PSP arrives, it causes an increment in the resting state potential of the neuron.

5. If many PSPs sum up, the resting state potential of the neuron can locally reach a certain threshold (approximately -55 mV), which causes the neuron to spike. This is called an action potential (AP), see Figure 2.6.

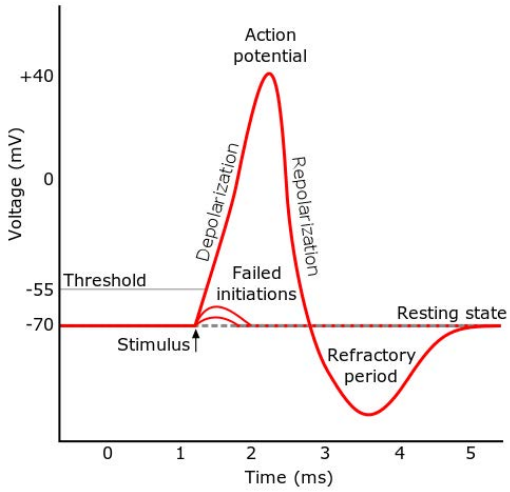


Figure 2.6: Approximate plot of a typical action potential (Source wikipedia.org).

2.3 How the neuron activity produces the electroencephalographic (EEG) signals

The current sources and the electromagnetic field generated by a single PSP or AP is not strong enough to be detected on the scalp. In order to have measurable signals, these tiny fields need to sum up. Action potentials have a temporal duration close to the millisecond making them hard to synchronize in order to sum up [Gramfort, 2009]. On the contrary, PSPs have a temporal

duration around 10 ms. Therefore, even though APs are much larger in amplitud than PSPs, it is accepted that the PSP are the generators of the scalp fields usually recorded in EEG [M. Hämäläinen et al., 1993]. If the dendrites supporting PSPs are randomly oriented or radially oriented on a complete spherical surface (or small closed surface), no net electromagnetic field can be detected outside the immediate vicinity of the active neurons. Pyramidal neurons are brain cells whose dendrites are perpendicular to the brain cortex (see Figure 2.7). Due to their uniform spatial organization, they are the only neurons that can generate a net current dipole over a piece of cortical surface, whose field is detectable on the scalp [Phillips, 2000]. According to [Murakami & Okada, 2006], it is necessary to add the field of approximately 10^4 pyramidal neurons in order to produce a field amplitude that is detectable on the scalp. Thus, the EEG measurements corresponds to the activity of one or several assemblies of neurons. The area of a neuron assembly is small compared to the distance to the observation point (the EEG sensors). Therefore, the electromagnetic fields produced by an active neuron assembly at the sensor level is very similar to the field produced by a current dipole [Gramfort, 2009]. This simplified model is known as the Equivalent Current Dipole (ECD). These ECDs are also known by several names, such as brain current sources, brain sources or simply sources.

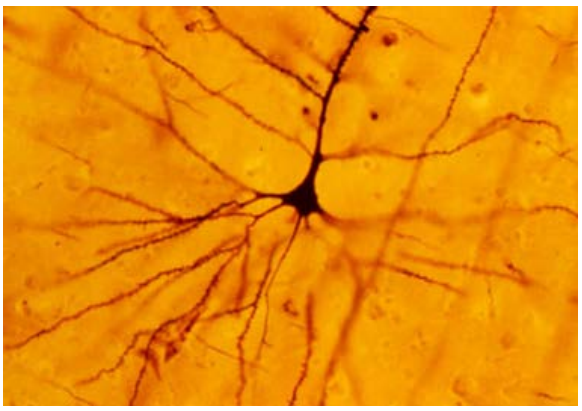


Figure 2.7: A human neocortical pyramidal neuron stained via Golgi technique (Source wikipedia.org).

2.4 EEG recording and measurement

2.4.1 A bit of history

The first EEG recordings date back to the first measurements by the English scientist Richard Caton in 1875. He used a galvanometer to observe electrical impulses from the surfaces of living brains in animal subjects [Collura, 1993]. Caton initially described the general positivity of the surface of the gray matter when measured in relation to deep structures. He also stated that, "The electric currents of the gray matter appear to have a relation to its function". For this work Caton is recognized as the discoverer of the EEG [Collura, 1993].

The first report of human EEG recording was made by the German physiologist and psychiatrist Hans Berger in 1929. Berger pointed out in his electrical brain recordings that there were regular rhythmic sequences of waves at about 10 cycles per second in the relaxed adult subject and that these were best seen with the eyes closed in the absence of stimulation or other mental activities such as imaging or problem solving. He called these waves alpha waves. He also observed smaller amplitude waves ranging in frequency from about 18 to 50 cycles per second. He called these waves beta waves. Berger called the entire electrical record of the brain's activity the Elektrenkephalogramm, abbreviated EEG, in keeping with the EKG acronym for the Electrokardiogramm [Patterson, 1973].

2.4.2 Methods and devices

In conventional EEG, the recording is obtained by placing electrodes on the scalp with a conductive gel or paste. Many systems typically use electrodes, each of which is attached to an individual wire. Some systems use caps or nets into which electrodes are embedded; this is particularly common when high-density arrays of electrodes are needed.

EEG systems consist of a number of delicate electrodes, a set of differential amplifiers (one for each electrode) followed by filters. The EEG recording electrodes and their proper function are crucial for acquiring high-quality data. Different types of electrodes are often used in the EEG recording systems, such as:

- disposable (gel-less, and pre-gelled types)
- reusable disc electrodes (gold, silver, stainless steel, or tin)
- headbands and electrode caps

- saline-based electrodes
- needle electrodes

For multichannel recordings, electrode caps are often used, see Figure 2.8. Commonly used scalp electrodes consist of Ag-AgCl disks, less than 3 mm in diameter, with long flexible leads that can be plugged into an amplifier. Needle electrodes are those that have to be implanted under the skull with minimal invasive operations. The conversion from analogue to digital EEG is performed by means of multichannel analogue-to-digital converters (ADCs). The raw EEG signals have amplitudes of the order of μvolts and contain frequency components of up to 100 Hz. To retain the effective information, the signals have to be amplified before the ADC and filtered, either before or after the ADC, to reduce the noise and make the signals suitable for processing and visualization. The filters are designed in such a way that they do not to introduce any change or distortion to the signals. Highpass filters with a cut-off frequency of usually less than 0.5 Hz are used to remove the disturbing very low frequency components such as those of breathing. On the other hand, high-frequency noise is mitigated by using lowpass filters with a cut-off frequency of approximately 50-70 Hz. Notch filters with a null frequency of 50 Hz are often necessary to ensure perfect rejection of the strong 50 Hz power supply.

(a) 256-128 channels (b) 64 channels (c) 32 channels

Figure 2.8: Multichannel EEG systems.

2.4.3 Electrode positioning

The clinical EEG is commonly recorded using the International 10/20 system, which is a standardized system for electrode placement [Jasper,

1958]. This particular recording system employs 21 electrodes attached to the surface of the scalp at locations defined by certain anatomical reference points; the numbers 10 and 20 are percentages signifying relative distances between different electrode locations on the skull perimeter, see Figure 2.9. The spacing of electrodes with the 10/20 system is relatively sparse: the interelectrode distance is approximately 4.5 cm on a typical adult head.

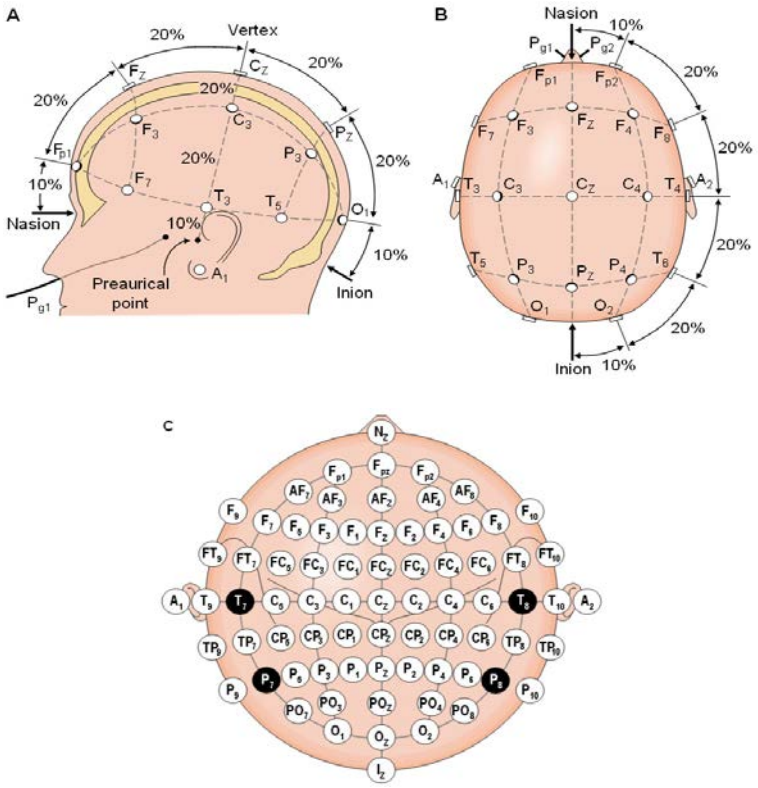


Figure 2.9: The international 10-20 system seen from (A) left and (B) above the head. A = Ear lobe, C = central, Pg = nasopharyngeal, P = parietal, F = frontal, Fp = frontal polar, O = occipital. (C) Location and nomenclature standardized by the American Electroencephalographic Society. (Adapted from [Sharbrough et al., 1991]).

Bipolar or unipolar electrodes can be used in the EEG measurement, see Figure 2.10. In the first method the potential difference between a pair of electrodes is measured. In the latter method the potential of each electrode is compared either with a neutral electrode or with the average of all electrodes

[Malmivuo & Plonsey, 1995].

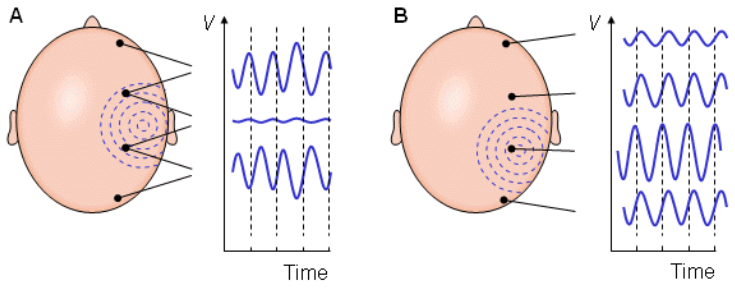


Figure 2.10: (A) Bipolar and (B) unipolar measurements. Note that the waveform of the EEG depends on the measurement location. (Adapted from [Malmivuo & Plonsey, 1995]).

2.4.4 Classification of EEG signals

The EEG signals are conventionally classified into five different frequency bands [Sörnmo & Laguna, 2005]:

- **Delta, <4 Hz:** it is typically encountered during deep sleep and has a large amplitude. It is usually not observed in the awake, normal adult, but is indicative of, e.g., cerebral damage or brain disease (encephalopathy).

- **Theta, 4-7 Hz:** it appears as consciousness slips towards drowsiness. Theta waves have been associated with access to unconscious material, creative inspiration and deep meditation.

- **Alpha, 8-13 Hz:** it appears in the posterior half of the head and are usually found over the occipital region of the brain. They can be detected in all parts of posterior lobes of the brain. Alpha waves have been thought to indicate both a relaxed awareness without any attention or concentration. Most subjects produce some alpha waves with their eyes closed, which is why it has been claimed that it is nothing but a waiting or scanning pattern produced by the visual regions of the brain. It is reduced or eliminated by opening the eyes, by hearing unfamiliar sounds, by anxiety, or mental concentration or attention.

- **Beta, 14-30 Hz:** This is a fast wave with low amplitude, associated with an activated cortex. It can be observed, for instance, during certain sleep stages. The beta wave is mainly observed in the frontal and central regions of the scalp.

- **Gamma, >30 Hz:** it is related to a state of active information processing of the cortex. Using an electrode located over the sensorimotor area and connected to a high-sensitivity recording technique, the gamma rhythm can be observed during finger movements [Andrew & Pfurtscheller, 1996].

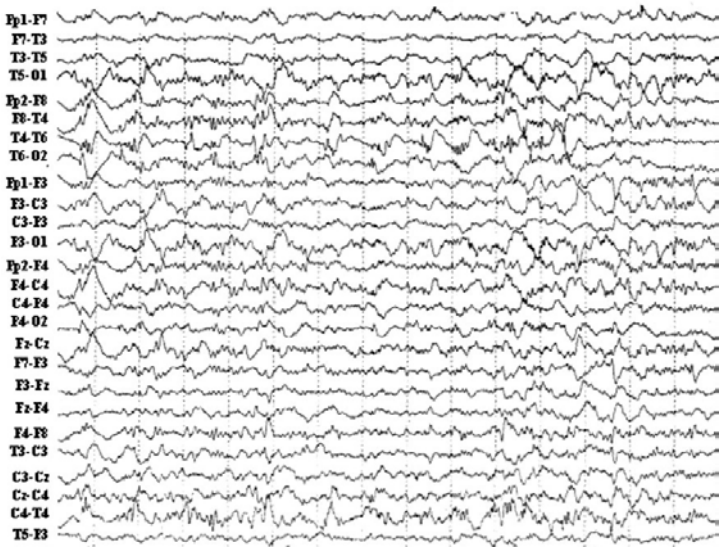


Figure 2.11: Sample EEG recordings. Electrodes have name (*e.g.*, F3, C3, *etc.*) function of their positions on the scalp (see Figure 2.9).

2.4.5 Event related potentials

Event related potentials (ERPs) are very small voltages generated in the brain structures in response to specific events or stimuli [Blackwood & Muir, 1990]. They are EEG changes that are time locked to sensory, motor or cognitive events that provide safe and noninvasive approach to study psychophysiological correlates of mental processes. They are thought to reflect the summed activity of postsynaptic potentials produced when a large

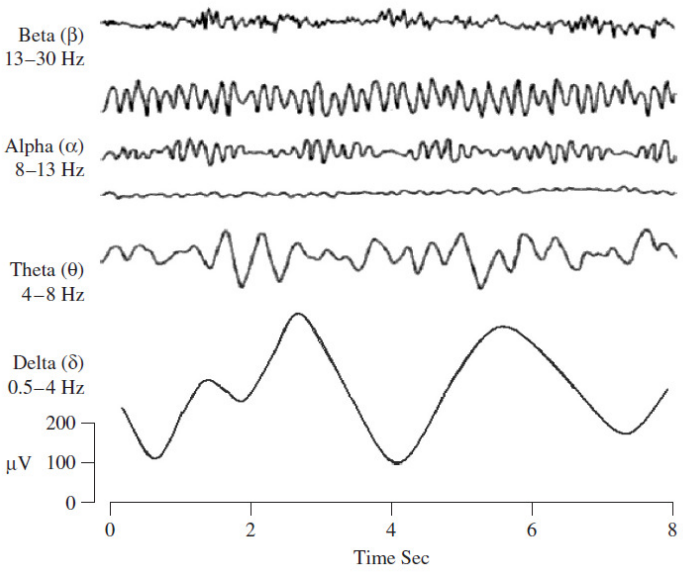


Figure 2.12: Four typical brain normal rhythms, from high to low frequencies. (Adapted from [Sanei & Chambers, 2008]).

number of similarly oriented cortical pyramidal neurons fire in synchrony while processing information [Peterson et al., 1995].

The EEG reflects thousands of simultaneously ongoing brain processes. This means that the brain response to a single stimulus or event of interest is not usually visible in the EEG recording of a single trial. Thus, to obtain an ERP, the experimenter must conduct many trials and average the results together, causing random brain activity to be averaged out and the relevant waveform to remain [Rugg & Coles, 1995].

ERP components are referred to by a letter (N/P) indicating polarity (negative/positive), followed by a number indicating either the latency in milliseconds or the component's ordinal position in the waveform, see Figure 2.13. For instance, a negative-going peak that is the first substantial peak in the waveform, and often occurs about 100 milliseconds after a stimulus is presented, is often called the N100 (indicating that its latency is 100 ms after the stimulus and that it is negative) or N1 (indicating that it is the first peak and is negative). The stated latencies for ERP components are often quite variable. For example, the P300 component may exhibit a peak anywhere between 250ms-700ms. [Luck, 2005].

Some examples of ERP components are the following:

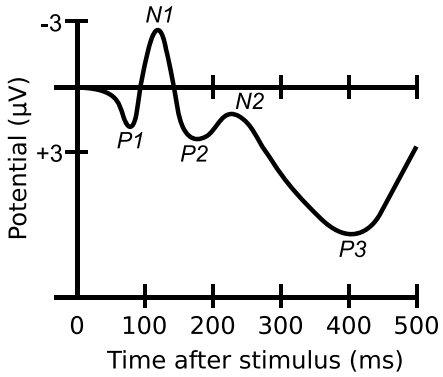


Figure 2.13: A waveform showing several ERP components, including the N100 and P300. Note that the ERP is plotted with negative voltages upward, a common, but not universal, practice in ERP research. (Source wikipedia.org)

- **N170:** it is a component of the event related potential that reflects the neural processing of faces. When potentials evoked by images of faces are compared to those elicited by other visual stimuli, the former show increased negativity 130-200 ms after stimulus presentation. This response is maximal over occipito-temporal electrode sites [Bentin et al., 1996].

- **P300:** it is usually interpreted as the speed of stimulus classification resulting from discrimination of one event from another [Sur & Sinha, 2009]. Shorter latencies indicate superior mental performance relative to longer latencies. P300 amplitude seems to reflect stimulus information such that greater attention produces larger P300 waves. A wide variety of paradigms have been used to elicit the P300, of which the "oddball" paradigm is the most utilized. In this paradigm, different stimuli are presented in a series such that one of them occurs relatively infrequently, that is the oddball. The subject is instructed to respond to the infrequent or target stimulus and not to the frequently presented or standard stimulus.

- **N400:** It is a negative wave first described in the context of semantic incongruity, 300-600 ms post-stimulus [Kutas & Hillyard, 1980]. N400 is inversely related to the expectancy of a given word to end a sentence.

- **Movement-related cortical potentials (MRCP):** MRCPs denote

a series of potentials that occur in close temporal relation with movement or movement-like activity. These may occur before, during or after the movement and they refer to the associated preparedness for movement in the cortex. [Deecke & Kornhuber, 1978] distinguished 4 components of the MRCPs: Bereitschafts potential (Readiness potential), Pre-motion positivity, Motor potential and Reafferent potential.

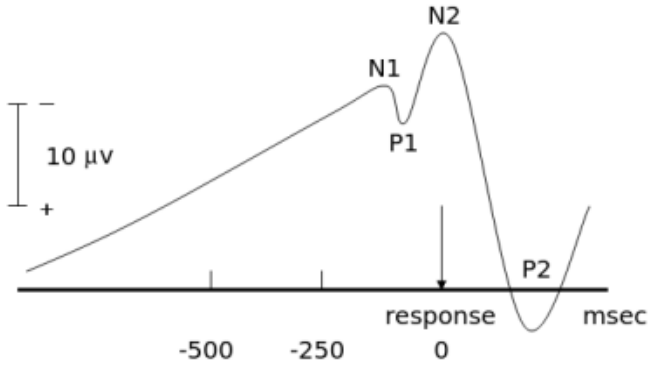


Figure 2.14: MRCP components: N1=Bereitschafts potential (Readiness potential), P1=Pre-motion positivity, N2=Motor potential, P2=Reafferent potential. (Adapted from wikipedia.org)

Chapter 3

EEG forward and inverse modelling

In this chapter, we introduce the EEG forward and inverse problems, and how they are related through the lead-field matrix. Next, we focus on the EEG inverse problem, which is the central component of the mobile brain scanner. Solving this problem will let us uncover the brain current sources whose electrical activity is producing the recorded EEG data. Additionally, we present some standard approaches to solve the EEG inverse problem. These approaches are based on the use of regularization functions, which help us to select, among the infinite possible solutions, the one that best fulfills some prescribed prior knowledge.

3.1 The EEG forward problem

The EEG forward problem consists in determining the EEG signals produced by the activity of a known distribution of current sources located inside the head, see Figure 3.1. This problem is well-posed, meaning that it has a unique solution, which is stable with respect to small perturbations in the data. To solve this problem, we first need to draw the attention towards the basis of EEG. According to [Plonsey & Heppner, 1967], no charge can be piled up in the conducting extracellular volume for the frequency range of the signals measured in the EEG. At one moment in time, all the fields are triggered by the active electric sources. Hence, no time delay effects are introduced. All fields and currents behave as if they were stationary at each instance. These conditions are also called quasi-static conditions. They are not static because the neural activity changes with time, but the changes

are slow compared to the propagation effects [Hallez et al., 2007].

Figure 3.1: Forward problem: given the position and the electrical activity of the sources, estimate the resulting EEG signals.

Under the quasi-static approximation to Maxwell's equations, the electric fields and magnetic fields are decoupled in the sense that the electric fields can be calculated or measured as if the magnetic fields do not exist and viceversa [Nunez & Srinivasan, 2006]. Thus, at time t, we can express the relation between the observed EEG vector $\mathbf{y}(t) \in \mathbb{R}^{M\times 1}$ and the current sources $\mathbf{s}(t) \in \mathbb{R}^{N\times 1}$ as a linear instantaneous form in the sources [Ou et al., 2009]:

$$\mathbf{y}(t) = \mathbf{A}\mathbf{s}(t) + \mathbf{e}(t) \tag{3.1}$$

where $\mathbf{e}(t) \in \mathbb{R}^{M\times 1}$ is the noise vector and M and N denote, respectively, the number of EEG electrodes and the number of sources. The relationship between the sensors and the current sources is given by the lead field matrix (forward model) $\mathbf{A} \in \mathbb{R}^{M\times N}$: the component $\mathbf{A}_{ij}$ denotes how the j-th source influences the measurement obtained by the i-th electrode. We can express the former model for all time instants $\{t_1, t_2, \ldots, t_T\}$ corresponding to some observation time window as follows:

$$\mathbf{Y} = \mathbf{A}\mathbf{S} + \mathbf{E} \tag{3.2}$$

where $\mathbf{Y} = [\mathbf{y}(t_1), \mathbf{y}(t_2), \ldots, \mathbf{y}(t_T)] \in \mathbb{R}^{M\times T}$ is the EEG measurements matrix, $\mathbf{S} = [\mathbf{s}(t_1), \mathbf{s}(t_2), \ldots, \mathbf{s}(t_T)] \in \mathbb{R}^{N\times T}$ is the sources matrix and $\mathbf{E} \in \mathbb{R}^{M\times T}$ is the noise matrix.

Taking into account the orientation of the pyramidal neurons mentioned in Section 2.3, we can now restrict the orientation of the sources to be perpendicular to the brain cortex. Throughout the rest of this thesis we will apply this belief. In order to determine the resulting EEG signals $\mathbf{Y}$,

first we need to compute the forward fields associated to each source. To accomplish this, we need to know a model of the head geometry, the EEG sensor positions, the conductivity values for each one of the tissues involved in the head model, and use all of these parameters to solve the corresponding quasi-static Maxwell's equations. For a more detailed overview on solvig the EEG forward problem see [Mosher et al., 1999; Hallez et al., 2007; Gramfort, 2009] and references therein.

3.1.1 Head Models

To describe the lead-field matrix **A** associated with current sources within the brain, we require a volume conductor model of the head. Modelling the conductivity layers of the head is a key requirement as the layers (tissues) attenuate the EEG signals differently due to their varying conductivity values [Stahlhut, 2011].

Spherical head models

The first volume conductor model of the human head consisted of a homogeneous sphere [Frank, 2004]. However it was soon noticed that the skull tissue had a conductivity which was significantly lower than the conductivity of scalp and brain tissue. Therefore, the volume conductor model of the head needed further refinement, and a three-shell concentric spherical head model was introduced, see Figure 3.2. In this model, the inner sphere represents the brain, the intermediate layer represents the skull and the outer layer represents the scalp [Hallez et al., 2007].

Realistic head models

The three-shell head model is a very simplified model of the human head. In reality the head is not spherical, besides, it is highly anisotropic and inhomogeneous. Therefore, more realistic head models are required, to obtain a more accurate solution to the forward problem. These head models use high resolution structural Magnetic Resonance Images (sMRI) to extract anatomical information, see Figure 3.3.

The different approaches that take this anatomical information into account to compute numerical solutions of the EEG forward problem are the following: Boundary Element Method (BEM), Finite Element Method (FEM) and the Finite Difference Method (FDM).

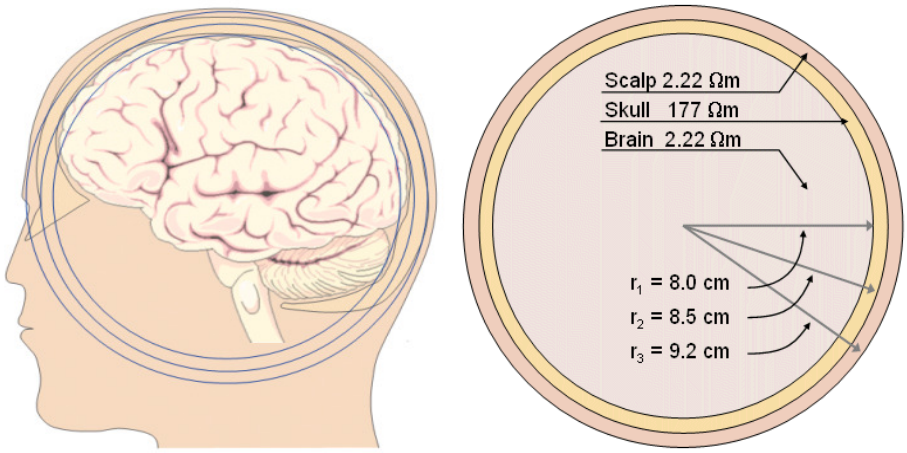


Figure 3.2: A spherical model with three layers. From inside to outside: brain cortex, skull and scalp. (Adapted from [Malmivuo & Plonsey, 1995]).

Boundary Element Method

The Boundary Element Method (BEM) is a numerical technique for calculating the surface potentials generated by current sources located in a piecewise homogeneous volume conductor. This method is capable of providing a solution to a volume problem by calculating the potential values (induced by a given current source) at the interfaces and boundary of the volume. In practice, a head model is built from several surfaces, each encapsulating a particular tissue. Typically, head models consist of 3 surfaces: brain-skull interface, skull-scalp interface and the outer surface. The regions between the interfaces are assumed to be homogeneous and isotropic conducting. To obtain a solution in such a piecewise homogenous volume, each interface is tesselated with small boundary elements [Hallez et al., 2007].

Finite Element Method

The Finite Element Method (FEM) is a technique, in which the entire 3D volume conductor is digitized into small elements, typically tetrahedrons. The most appealing factor of the FEM is the possibility of modeling tissues as anisotropic [C. Wolters et al., 2002; C. H. Wolters et al., 2004]. A drawback of the FEM is that it is normally regarded as being quite time consuming, due to the greatly increased computation complexity involved

Figure 3.3: Structural Magnetic Resonance Image, sagittal view. (Source http://www2.fmrib.ox.ac.uk).

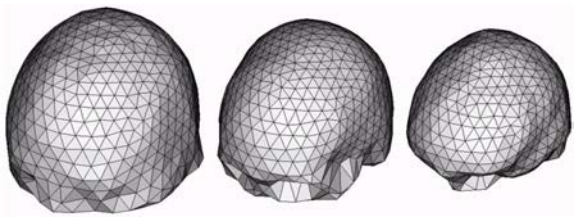


Figure 3.4: Example mesh of the human head used in BEM. From left to right: scalp, outer skull and inner skull. (Adapted from [Im & He, 2006]).

by tesselating a volumen instead of a surface as in the BEM.

Finite Difference Method

An alternative realistic head model can be obtained by the Finite Difference Method (FDM), which uses a cubic grid to discretize the volume conductor. The main drawback of the FDM method for EEG forward modelling is that, due to the cubic grid, the complex interfaces between brain structures and thin layers cannot be precisely modeled [Gramfort, 2009].

In summary, the difference between BEM, FEM, and FDM is the domain in which they are calculated. BEM solves the forward problem by boundaries between homogeneous isotropic compartments, whereas FEM and FDM solve the forward problem with an entire volume. Consequently, FEM and FDM easily lead to larger number of computational points than BEM. Typically, values for BEM are in the order of 5000 to 25000 computational points whereas the computational points involved in FEM/FDM solutions

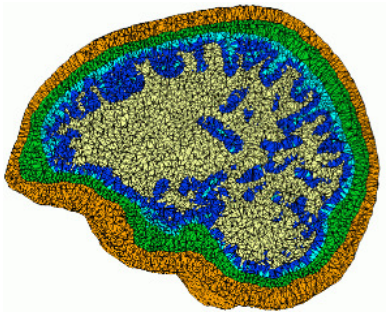


Figure 3.5: Example mesh of the human head used in FEM. From inside to outside: white matter, gray matter, cerebrospinal fluid, skull and scalp. (Adapted from [C. Wolters et al., 2006]).

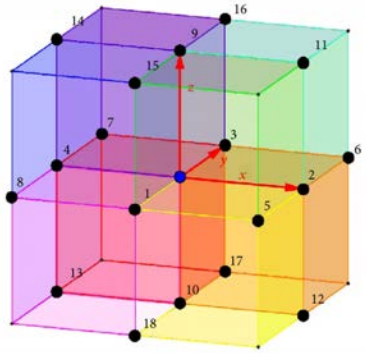


Figure 3.6: Example of the stencil used in FDM. A typical node in a FDM grid with its neighbours. (Adapted from [Turovets et al., 2014]).

are in the range of 10^4 to 10^5 [Stahlhut, 2011]. Taking into account the restricted resources of a mobile device, throughout this thesis we will use a three-shell spherical head model.

3.2 The EEG inverse problem

Given the lead-field matrix $\mathbf{A}$ and the EEG measurements matrix $\mathbf{Y}$, the EEG inverse problem consists in estimating the sources matrix $\mathbf{S}$ by solving the linear estimation problem (3.2), see Figure 3.7. Before we address this problem, let us examine some important features related to it. A standand EEG neuroheadset has approximately 10-200 EEG sensors,

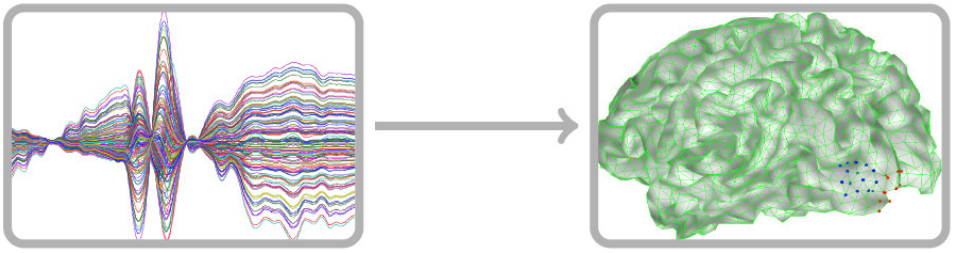


Figure 3.7: Inverse problem: Given the EEG signals, find the position and the electrical activity of the sources that have produced the recorded data.

whereas the typical number of possible source locations is approximately 10^3-10^5 (depending on the choice of the head model and the spatial resolution of interest). Thus, the small number of EEG sensors available to capture the electrical activity produced by thousands of brain electrical current sources, plus the low spatial resolution of the EEG, caused by the low conductivity of the skull, which blurs the electrical potential on the scalp, imply that the EEG inverse problem is an underdetermined ill-posed problem with an infinite number of solutions. From the electromagnetic point of view, the nonuniqueness of the EEG inverse problem can be derived from the fact that there are so-called electrically silent sources, which produce no electric field outside the head. Such a current source can always be added to a solution of the inverse problem without affecting the field outside the head [Sarvas, 1987]. Besides the ill-posed nature of the EEG inverse problem, another serious concern is its numerical instability often referred to as being ill-conditioned, which will be discussed in the following section.

3.2.1 Numerical sensitivity of the EEG inverse problem

Without loss of generality, let us analyze the numerical sensitivity of (3.2) in the noiseless case:

$$\mathbf{Y} = \mathbf{AS} = (\mathbf{U\Sigma V}^\top)\mathbf{Y} \tag{3.3}$$

where $\mathbf{A} = \mathbf{U\Sigma V}^\top$ is the Singular Value Decomposition (SVD) of $\mathbf{A}$. Using the pseudoinverse of the lead-field matrix , $\mathbf{A}^\dagger = \mathbf{V\Sigma}^\dagger\mathbf{U}^\top$, we can estimate the sources matrix as follows:

$$\mathbf{S} = \mathbf{A}^\dagger\mathbf{Y} = (\mathbf{V\Sigma}^\dagger\mathbf{U}^\top)\mathbf{Y} \tag{3.4}$$

where $\mathbf{\Sigma}^\dagger = \text{diag}(1/\sigma_1, 1/\sigma_2, \ldots, 1/\sigma_p)$. To examine how perturbations in the EEG measurements matrix $\mathbf{Y}$ affect the sources matrix $\mathbf{S}$, let us expand

(3.4) as follows:

$$\mathbf{S}_j = (\mathbf{V}\boldsymbol{\Sigma}^{\dagger}\mathbf{U}^{\top})\mathbf{Y}_j = \sum_{i=1}^{p} \frac{\mathbf{U}_i^{\top}\mathbf{Y}_j}{\sigma_i}\mathbf{V}_i \tag{3.5}$$

where $\mathbf{S}_j$ and $\mathbf{Y}_j$ denote the j-th column of $\mathbf{S}$ and $\mathbf{Y}$, respectively. This expansion shows that small changes in $\mathbf{Y}$ can induce relatively large changes in $\mathbf{S}$ if σ_i is small. To formalize this assertion, let us compute the relative error in the values of the sources at time t_j ($\delta\mathbf{S}_j$) caused by a perturbation ocurred at the same time in the EEG measurements ($\delta\mathbf{Y}_j$):

$$\mathbf{A}\delta\mathbf{S}_j = \delta\mathbf{Y}_j \quad \Rightarrow \quad \|\delta\mathbf{S}_j\|_p \leq |||\mathbf{A}^{\dagger}|||_p \, \|\delta\mathbf{Y}_j\|_p \tag{3.6}$$

$$\mathbf{A}\mathbf{S}_j = \mathbf{Y}_j \quad \Rightarrow \quad \frac{1}{\|\mathbf{S}_j\|_p} \leq \frac{|||\mathbf{A}|||_p}{\|\mathbf{Y}_j\|_p} \tag{3.7}$$

where $\|\cdot\|_p$ and $|||\cdot|||_p$ denote, respectively, the ℓ_p-vector norm and the ℓ_p-matrix norm [Horn & Johnson, 1990]. Combining (3.6) and (3.7), we get:

$$\frac{\|\delta\mathbf{S}_j\|_p}{\|\mathbf{S}_j\|_p} \quad \leq \quad \kappa_p(\mathbf{A})\frac{\|\delta\mathbf{Y}_j\|_p}{\|\mathbf{Y}_j\|_p} \tag{3.8}$$

The constant $\kappa_p(\mathbf{A}) = |||\mathbf{A}|||_p|||\mathbf{A}^{\dagger}|||_p$ is called the condition number of the system (matrix). Note that $\kappa_p(\cdot)$ depends on the underlying norm and subscripts are used accordingly. Nevertheless, it is usual to compute the condition number using $p = 2$:

$$\kappa(\mathbf{A}) = \kappa_2(\mathbf{A}) = |||\mathbf{A}|||_2 \, |||\mathbf{A}^{\dagger}|||_2 = \frac{\sigma_{\max}(\mathbf{A})}{\sigma_{\min}(\mathbf{A})} \tag{3.9}$$

where $|||\mathbf{A}|||_2 = \sigma_{\max}(\mathbf{A})$ denotes the spectral norm of $\mathbf{A}$.

The inequality (3.8) shows that the relative error is determined by the condition number $\kappa(\mathbf{A})$. Systems with relatively large condition number are said to be ill-conditioned. These kind of systems are numerically unstable, which means that small variations in the input ($\mathbf{Y}$) lead to large variations in the solution ($\mathbf{S}$). In Figure 3.8, we can see the typical behavior of the singular values associated to a lead-field matrix $\mathbf{A}$: they decay gradually to zero. Therefore, $\kappa(\mathbf{A})$ is large, which makes the EEG inverse problem ill-conditioned. This behaviour in the singular values is common in problems that result from the discretization of continuous ill-posed problem. In the case of the EEG inverse problem, the continuous ill-posed problem involved

is the following Fredholm integral equation of the first kind [Trujillo-Barreto et al., 2004; Wing, 1991]:

$$\mathbf{Y}(\overrightarrow{\mathbf{r}}_i, t) = \int \mathbf{A}(\overrightarrow{\mathbf{r}}_i, \overrightarrow{\mathbf{r}}_j)\mathbf{S}(\overrightarrow{\mathbf{r}}_j, t)d^3\overrightarrow{\mathbf{r}}_j \tag{3.10}$$

It is well-known that problems arisen from the discretization of Fredholm integral equations of the first kind lead to ill-conditioned problems [Hansen, 1998]. The overall behavior of the singular values is strongly connected with the properties of the lead-field kernel $\mathbf{A}(\overrightarrow{\mathbf{r}}_i, \overrightarrow{\mathbf{r}}_j)$ [Hansen, 1998]: the "smoother" the kernel, the faster the singular values decay to zero (where "smoothness" is measured by the number of continuous derivatives of $\mathbf{A}$). From the physical point of view, this smoothness may be explained by the blurring effect induced by the low conductivity of the skull.

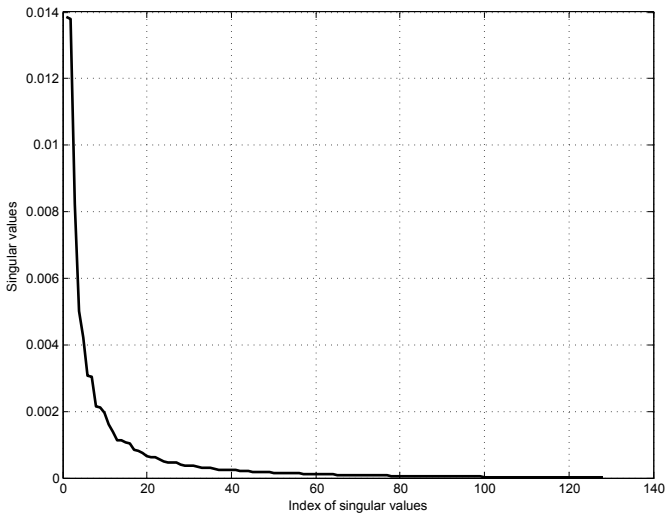


Figure 3.8: Singular values of a lead-field matrix $\mathbf{A}$ associated to a three-shell spherical head model.

A classical approach to cope with the ill-posedness and nonuniqueness of the EEG inverse problems is to use regularization methods, which involve the replacement of the original ill-posed problem with a nearby well-posed problem whose solution approximates the required solution.

3.3 Regularization of the EEG inverse problem

In order to regularize the EEG inverse problem, we need to use a regularization function, which helps us to select, among the infinite solutions, the one that best fulfills some prescribed prior knowledge. As prior knowledge, we can use a mathematical constrain (minimum norm solution) or use some known anatomical, physiological and functional feature of the human brain, such as the smoothness and focality of the electromagnetic fields generated and propagated within the volume conductor media (brain cortex, skull and scalp), the dynamics of the electrical activity of the sources, the clusters formed by neighboring or functional related sources, among others [M. Hämäläinen et al., 1993; Menendez et al., 2004].

The regularized EEG inverse problem can be stated as follows:

$$\hat{\mathbf{S}} = \underset{\mathbf{S}}{\operatorname{argmin}} \left\{ \frac{1}{2} ||\mathbf{AS} - \mathbf{Y}||_F^2 + \lambda \Omega(\mathbf{S}) \, , \lambda > 0 \right\} \tag{3.11}$$

where $\frac{1}{2}||\mathbf{AS} - \mathbf{Y}||_F^2$ is the reconstruction error ($\| \ \|_F$ denotes the Frobenius norm), $\Omega(\mathbf{S})$ is the regularization function (also called penalty term) and λ is the regularization parameter, which provides a tradeoff between the minimization of the reconstruction error and the fulfillment of the prior knowledge encoded by $\Omega(\mathbf{S})$. The regularized EEG inverse problem can be interpreted as follows: among the infinite possible solutions, select the one that best fulfills the prior knowledge encoded by the regularization function $\Omega(\mathbf{S})$ and, at the same time, has the minimum reconstruction error. If the regularization function $\Omega(\mathbf{S})$ is a convex function, then (3.11) is a convex optimization problem [S. P. Boyd & Vandenberghe, 2004], which has several advantages, including: it has a global (unique) minimum, there are several reliable and efficient numerical methods for solving it [S. Wright & Nocedal, 1999] and there exist several numerical and analytic tools for analyzing its mathematical properties [Rockafellar, 1997; Borwein et al., 2006; D. P. Bertsekas, 2009].

In the following sections, we will visit some of the most commonly used regularization functions in the EEG inverse problem.

3.3.1 Minimum norm solution

The minimum norm regularization, known as Minimum Norm Estimate (MNE) in the EEG community [M. S. Hämäläinen & Ilmoniemi, 1994], is

based on the squared Frobenius norm:

$$\hat{\mathbf{S}} = \underset{\mathbf{S}}{\operatorname{argmin}} \left\{ \frac{1}{2} ||\mathbf{AS} - \mathbf{Y}||_F^2 + \frac{\lambda}{2} \|\mathbf{S}\|_F^2 \right\} \tag{3.12}$$

This regularization function usually induces a solution that tends to be spread over a considerable part of brain.

(3.12) admits a closed-form solution, which can be obtained as follows:

$$\begin{aligned} &\frac{\partial}{\partial \mathbf{S}} \left[\frac{1}{2} \|\mathbf{A}\hat{\mathbf{S}} - \mathbf{Y}\|_F^2 + \frac{\lambda}{2} \|\hat{\mathbf{S}}\|_F^2 \right] = 0 \\ &\Rightarrow \mathbf{A}^\top (\mathbf{A}\hat{\mathbf{S}} - \mathbf{Y}) + \lambda \hat{\mathbf{S}} = 0 \\ &\Rightarrow \hat{\mathbf{S}} = (\mathbf{A}^\top \mathbf{A} + \lambda \mathbf{I}_N)^{-1} \mathbf{A}^\top \mathbf{Y} \end{aligned} \tag{3.13}$$

The fact that the inverse solution is given by a simple matrix multiplication makes the MNE an attractive approach. Nevertheless, it can happen that computing the inverse matrix is intractable in practice due to its computational complexity (the number of operations required to compute it) is $O(N^3)$. Besides, it would require to store a matrix of dimension $N \times N$. For instance, let us assume we want to use (3.13) to solve an EEG inverse problem involving $N = 20000$ sources and $M = 100$ EEG electrodes. This would involve 8×10^{12} floating point operations and the storage of a matrix of dimension 20000×20000, which would require ≈ 3.2 GB of RAM. In order to avoid these problems, we can use the Sherman-Morrison-Woodbury formula to compute the inverse solution as follows:

$$\hat{\mathbf{S}} = \mathbf{A}^\top (\mathbf{A}\mathbf{A}^\top + \lambda \mathbf{I}_M)^{-1} \mathbf{Y}$$

In this case, we would need to compute the inverse of a much smaller matrix of dimension 100×100. This would involve 10^6 floating point operations and the storage of this small matrix would require $\approx$ 80KB of RAM.

3.3.2 Low Resolution Electromagnetic Tomography: LORETA

The LORETA regularization approach [R. D. Pascual-Marqui et al., 1994] assumes that neighboring sources are simultaneously and synchronously activated. This basic assumption rests on evidence from single cell recording in the brain that demostrates strong syncronization of adjacent neurons. LORETA leads to the following regularized EEG inverse problem:

$$\hat{\mathbf{S}} = \underset{\mathbf{S}}{\operatorname{argmin}} \left\{ \frac{1}{2} ||\mathbf{AS} - \mathbf{Y}||_F^2 + \frac{\lambda}{2} \|\mathbf{BWS}\|_F^2 \right\} \tag{3.14}$$

where $\mathbf{B}$ is the discrete Laplacian operator of the cortical mesh (to induce the spatial smoothness) [R. Pascual-Marqui, 1999] and $\mathbf{W} = \text{diag}(\|\mathbf{A}_1\|, \|\mathbf{A}_2\|, \ldots, \|\mathbf{A}_N\|)$ is a weighting matrix that takes into account the strength of the forward field of each source to avoid biased solutions toward the surface of the cortex. (3.14) admits the following closed-form solution:

$$\hat{\mathbf{S}} = (\mathbf{W}\mathbf{B}^\top\mathbf{B}\mathbf{W})^{-1}\mathbf{A}^\top \left(\mathbf{A}(\mathbf{W}\mathbf{B}^\top\mathbf{B}\mathbf{W})^{-1}\mathbf{A}^\top + \lambda\mathbf{I}_M\right)^{-1}\mathbf{Y} \tag{3.15}$$

3.3.3 Lasso regularization

The Lasso regularization [Tibshirani, 1996] (known as basis pursuit [Chen et al., 1999] in the signal processing community and Minimum Current Estimate (MCE) [Uutela et al., 1999] in the context of EEG inverse problems), uses the ℓ_1-norm to enconde the following prior knowledge: during a particular cognitive task, only the sources related with the brain area involved in such a task will be active, therefore, it is expected a solution with few nonzero components. Formally, the Lasso regularization can be expressed as follows:

$$\hat{\mathbf{S}} = \underset{\mathbf{S}}{\text{argmin}} \left\{\frac{1}{2}\|\mathbf{A}\mathbf{S} - \mathbf{Y}\|_F^2 + \lambda\|\mathbf{S}\|_1\right\} \tag{3.16}$$

where $\|\mathbf{S}\|_1 = \sum_{i=1}^{N}\sum_{j=1}^{T}|s_{ij}|$. Regularizing by the ℓ_1-norm is known to induce sparsity in the sense that, a number of components of $\hat{\mathbf{S}}$, depending on the strength of the regularization parameter, will be exactly zero [Friedman et al., 2001]. A geometrical interpretation of this fact can be observed by casting (3.16) into a constrained optimization problem:

$$\begin{aligned} \hat{\mathbf{S}} \quad &= \quad \underset{\mathbf{S}}{\text{argmin}} \left\{\frac{1}{2}\|\mathbf{A}\mathbf{S} - \mathbf{Y}\|_F^2\right\} \\ &\text{s.t} \quad \|\mathbf{S}\|_1 \leq \mu \end{aligned} \tag{3.17}$$

for some $\mu > 0$ (the set of solutions of (3.16) and (3.17) are the same, as described by some value $\lambda(\mu)$ depending on μ [Borwein et al., 2006]). At optimality, the level set of $\frac{1}{2}\|\mathbf{A}\mathbf{S} - \mathbf{Y}\|_F^2$ is tangent to the set $\{\mathbf{S} \in \mathbb{R}^{N\times T};\ \|\mathbf{S}\|_1 \leq \mu\}$, hence, as we can see in Figure 3.9, the geometry of the norm's ball is directly related to the properties of the solutions $\hat{\mathbf{S}}$. If $\Omega(\mathbf{S})$ is taken to be the squared Frobenius norm, then the resulting ball is the standard, isotropic, "round" ball that does not favor any specific direction of the space. On the other hand, when $\Omega(\mathbf{S})$ is the ℓ_1-norm, its

corresponding ball presents a diamond-shaped pattern in two dimensions, and a pyramid in three dimensions. In particular, this ball is anisotropic and exhibits some singular points due to the extra nonsmoothness of the ℓ_1-norm. Moreover, these singular points are located along the axis, so that if the level set happens to be tangent at one of those points, sparse solutions are obtained (see Figure 3.10) [Bach et al., 2011].

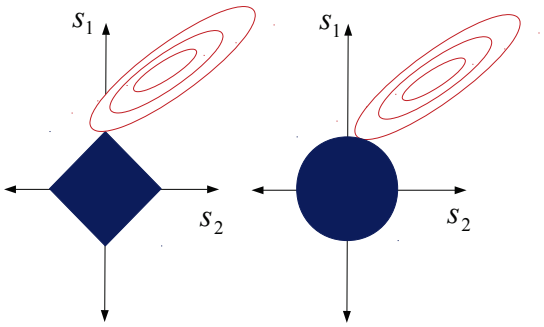


Figure 3.9: In blue, balls for the ℓ_1-norm and ℓ_2-norm. In red, some level sets of the squared loss function are plotted. At optimality, the level sets are tangent to the blue balls.

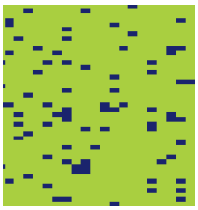


Figure 3.10: Sparsity pattern induced by the ℓ_1-norm (nonzero components are shown in blue).

3.3.4 Group Lasso regularization

In some situations, the components of the sources matrix $\mathbf{S}$ can be naturally partitioned in groups. When this happens, we would like to select or

remove simultaneously all the components forming a group. The Group Lasso regularization [Turlach et al., 2005; Yuan & Lin, 2006] uses the $\ell_{2,1}$-norm to exploit this group structure. For instance, if the components of the sources matrix are partitioned by rows, and we would like to select or remove simultaneously all the components forming a row, the following Group Lasso regularization function can be used to encode this behavior:

$$\|\mathbf{S}\|_{2,1} = \sum_{i=1}^{N} \|\mathbf{S}(i,:)\|_2 \tag{3.18}$$

where $\mathbf{S}(i,:)$ denotes the i-th row of the sources matrix $\mathbf{S}$. It behaves like an ℓ_1-norm on the vector $(\|\mathbf{S}(i,:)\|_2)_{i=1}^{N}$, and therefore, it induces group sparsity: each $\|\mathbf{S}(i,:)\|_2$, and equivalently each $\mathbf{S}(i,:)$, is encouraged to be set to zero [Bach et al., 2011]. These type of solutions, which involve few nonzero components forming a well-defined pattern, are usually called structured sparse solutions (see Figure 3.11). From the neurophysiological point of view, the Group Lasso regularization can be used to extend the prior knowledge introduced by Lasso: it is expected a solution with few nonzero components, but in addition, these nonzero components will be grouped by rows, that is, it is expected a solution with few nonzero rows (once a source is active, it will remain active during the observation time window). Formally, the Group Lasso regularization can be expressed as follows:

$$\hat{\mathbf{S}} = \underset{\mathbf{S}}{\operatorname{argmin}} \left\{ \frac{1}{2}\|\mathbf{AS} - \mathbf{Y}\|_F^2 + \lambda\|\mathbf{S}\|_{2,1} \right\} \tag{3.19}$$

The $\ell_{2,1}$-norm penalty has also been used in regularization approaches that involve the decomposition of the sources matrix $\mathbf{S}$ as linear combinations of multiple basis functions, $\mathbf{S} = \Theta\Phi^\top$, where $\Phi^\top \in \mathbb{R}^{P\times T}$ is a basis matrix and $\Theta \in \mathbb{R}^{N\times P}$ is a matrix that contains the coefficients of the decomposition. In theses cases, the basis matrix is usually fixed whereas the $\ell_{2,1}$-norm is used to induce structured sparse solution by regularizing the coefficient matrix as follows:

$$\hat{\mathbf{\Theta}} = \underset{\mathbf{\Theta}}{\operatorname{argmin}} \left\{ \frac{1}{2}\|\mathbf{A\Theta\Phi}^\top - \mathbf{Y}\|_F^2 + \lambda\|\mathbf{\Theta}\|_{2,1} \right\} \tag{3.20}$$

These approaches often differ in the basis matrix used in the decomposition. For instance, [Ou et al., 2009] uses temporal basis functions obtained from the SVD of $\mathbf{Y}$, [Gramfort et al., 2011, 2013] uses time frequency Gabor dictionaries and [Haufe et al., 2011] uses spatial basis Gaussian functions.

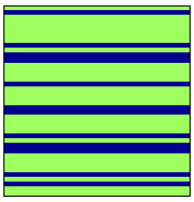


Figure 3.11: Sparsity pattern induced by the $\ell_{2,1}$-norm defined in (3.18).

3.3.5 Sparse Group Lasso regularization

In some situations, it is desirable to select or remove simultaneously all the components forming a group and, at the same time, induce sparsity inside the selected groups. The Sparse Group Lasso regularization [Friedman et al., 2010; Sprechmann et al., 2011] uses the $(\ell_{2,1} + \ell_1)$-norm to exploit this structure:

$$\|\mathbf{S}\|_{2,1} + \|\mathbf{S}\|_1 = \sum_{i=1}^{N} \left(\|\mathbf{S}(i,:)\|_2 + \sum_{j=1}^{T} |s_{ij}| \right) \tag{3.21}$$

In the EEG inverse problem, we can use this regularization function to introduce the following prior knowledge: when the observation time window is wide (with respect to the temporal dynamics of the neural signals), it is expected that the sources change their state (on-off) one or several times (depending on the amplitude of the observation window). Therefore, in this case, it is expected a solution with few nonzero rows, with sparsity inside each one of them (see Figure 3.12). Formally, the Sparse Group Lasso regularization can be expressed as follows:

$$\hat{\mathbf{S}} = \underset{\mathbf{S}}{\operatorname{argmin}} \left\{ \frac{1}{2}\|\mathbf{AS} - \mathbf{Y}\|_F^2 + \lambda \left(\rho \|\mathbf{S}\|_{2,1} + (1-\rho)\|\mathbf{S}\|_1 \right), \ \rho \in (0,1) \right\} \tag{3.22}$$

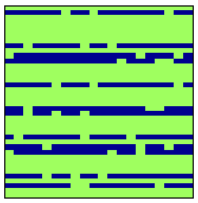


Figure 3.12: Sparsity pattern induced by the $(\ell_{2,1} + \ell_1)$-norm defined in (3.21).

3.3.6 Trace norm regularization

The regularization functions that we have seen so far are useful to introduce prior knowledge related with the sparsity pattern of the sources matrix $\mathbf{S}$, but they do not take into account possible functional dependencies among the active sources. This kind of prior knowledge could be helpful to describe the following scenario: it is expected that the main active sources induce an electrical activity in their corresponding neighboring sources. Therefore, there is a dependency between the induced activity and the precursor activity, which could be modeled, for instance, as a linear relationship. Hence, the expected solution will have a low rank. This prior knowledge can be exploited using the Trace norm (Nuclear norm) regularization [Argyriou et al., 2007; Abernethy et al., 2009], which is defined as the sum of the singular values:

$$\|\mathbf{S}\|_* = \sum_{i=1}^{q} \sigma_i(\mathbf{S}) \tag{3.23}$$

where $q = \min\{N, T\}$ and $\sigma_i(\mathbf{S})$ denotes the i-th singular value of $\mathbf{S}$. This norm is the convex relaxation of the rank function (the prior knowledge previously mentioned could also be exploited using the rank function directly, but this would transform the optimization problem into a combinatorial NP-hard problem, since it includes the cardinality minimization as a special case [Zhao, 2012; Recht et al., 2010; Natarajan, 1995]). Formally, the Trace norm regularization is defined as follows:

$$\hat{\mathbf{S}} = \underset{\mathbf{S}}{\operatorname{argmin}} \left\{ \frac{1}{2}\|\mathbf{AS} - \mathbf{Y}\|_F^2 + \lambda\|\mathbf{S}\|_* \right\} \tag{3.24}$$

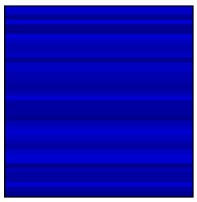


Figure 3.13: Dense pattern induced by the ℓ_*-norm defined in (3.23). This norm induces a dense low rank solution. Components with the same blue intensity are linear dependent.

As we have seen in Sections 3.3.1 and 3.3.2, when we use the MNE and LORETA regularization functions, the regularized EEG inverse problem

(3.11) can be solved in closed-form. However, this is not the case when we use the Lasso, Group Lasso, Sparse Group Lasso and Trace Norm regularization functions. One characteristic common to these regularization functions is that they are nonsmooth (nondifferentiable), unlike MNE and LORETA, both of which are smooth. In the following section, we will see how to solve the regularized EEG inverse problem when the regularization function involved is nonsmooth.

3.4 Solving the regularized nonsmooth EEG inverse problem

A standand approach to deal with the nonsmoothness introduced by the ℓ_1, $\ell_{2,1}$, $\ell_{2,1} + \ell_1$ and ℓ_* regularization functions is to reformulate the regularization problem as a second-order cone programming (SOCP) problem [Haufe et al., 2011] and use interior point-based solvers. However, interior point-based methods can not handle large scale problems, which is the case of large EEG inverse problems involving thousands of sources. Another approach is to try to solve the nonsmooth problem directly, using general nonsmooth optimization methods, for instance, the subgradient method [D. Bertsekas, 1999]. This method can be used if a subgradient of the objective function can be computed efficiently [Bach et al., 2011]. However, its convergence rate is, in practice, slow ($O(1/\sqrt{k})$), where k is the iteration counter. In this thesis, in order to tackle the regularized nonsmooth EEG inverse problem, we depart from these optimization methods and use instead efficient first-order nonsmooth optimization methods [Combettes & Wajs, 2005; Combettes & Pesquet, 2011; Beck & Teboulle, 2009]: forward-backward splitting methods. These methods are also called proximal splitting because the nonsmooth function is involved via its proximity operator, which is defined as follows:

Definition 3.4.1. *Proximity operator. The proximity operator [Moreau, 1965] corresponding to a convex function F is a mapping from $\mathbb{R}^N$ to itself and is defined as follows:*

$$\text{prox}_F(\mathbf{Z}) = \underset{\mathbf{X}}{\text{argmin}} \left\{ F(\mathbf{X}) + \frac{1}{2}\|\mathbf{X} - \mathbf{Z}\|^2 \right\} \tag{3.25}$$

where $\|\cdot\|$ denotes the Euclidean norm. Note that the proximity operator is well defined, because the above minimum exists and is unique.

Forward-backward splitting methods were first introduced in the EEG inverse problem by [Gramfort et al., 2011], where they used them to solve nonsmooth optimization problems resulting from the use of the $\ell_{2,1}$-norm penalty function. These methods have drawn increasing attention in the EEG, machine learning, and signal processing community, especially because of their convergence rates and their ability to deal with large problems [Combettes & Pesquet, 2011; Nesterov, 2007; S. J. Wright et al., 2009].

3.4.1 Forward-Backward splitting methods

Forward-backward splitting methods (also called proximal splitting methods) are specifically tailored to solve an optimization problem of the form

$$\underset{\mathbf{S}}{\text{minimize}}\ F(\mathbf{S}) + \Omega(\mathbf{S}) \tag{3.26}$$

where $F(\mathbf{S})$ is a smooth convex function, and $\Omega(\mathbf{S})$ is also a convex function, but nonsmooth (in the case of the regularized EEG inverse problem, $F(\mathbf{S}) = \frac{1}{2}\|\mathbf{AS} - \mathbf{Y}\|_F^2$, whereas $\Omega(\mathbf{S})$ is one of the nonsmooth regularization functions mentioned in the previous section). From convex analysis [D. Bertsekas, 1999], we know that $\mathbf{S}$ is a minimizer of (3.26) if and only if $0 \in \partial(F + \Omega)(\mathbf{S})$, where $\partial(F + \Omega)(\mathbf{S})$ denotes the subdifferential of $(F + \Omega)$ at $\mathbf{S}$. This implies the following [Combettes & Wajs, 2005]:

$$\begin{aligned} 0 \in \partial(F + \Omega)(\mathbf{S}) &\Leftrightarrow 0 \in \{\partial F(\mathbf{S}) + \partial\Omega(\mathbf{S})\} \\ &\Leftrightarrow -\nabla F(\mathbf{S}) \in \partial\Omega(\mathbf{S}) \\ &\Leftrightarrow -\gamma\nabla F(\mathbf{S}) \in \gamma\partial\Omega(\mathbf{S}) \\ &\Leftrightarrow (\mathbf{S} - \gamma\nabla F(\mathbf{S})) - \mathbf{S} \in \partial\gamma\Omega(\mathbf{S}) \end{aligned}$$

Taking into account the following subdifferential-proximity operator relationship, $\mathbf{X} \in \partial G(\mathbf{Z}) \Leftrightarrow \mathbf{Z} = \text{prox}_G(\mathbf{X} + \mathbf{Z})$[Micchelli et al., 2011], we get the following:

$$\mathbf{S} = \text{prox}_{\gamma\Omega}(\mathbf{S} - \gamma\nabla F(\mathbf{S})) \tag{3.27}$$

Equation (3.27) suggests that we can solve (3.26) using a fixed point iteration:

$$\mathbf{S}_{k+1} = \text{prox}_{\gamma\Omega}(\mathbf{S}_k - \gamma\nabla F(\mathbf{S}_k)) \tag{3.28}$$

In optimization, (3.28) is known as forward-backward splitting process [Combettes & Pesquet, 2011]. It consists of two steps: first, it performs a forward gradient descend step $\mathbf{S}_k^* = \mathbf{S}_k - \gamma\nabla F(\mathbf{S}_k)$ and then it performs a backward step $\mathbf{S}_{k+1} = \text{prox}_{\gamma\Omega}(\mathbf{S}_k^*)$.

From (3.28), we can see the importance of the proximity operator (associated to the nonsmooth regularization function $\gamma\Omega(\mathbf{S})$) with respect to the forward-backward splitting methods, since their main step is to calculate it. If we would have a closed-form expression for such proximity operator or if we could approximate it efficiently (with the approximation errors decreasing at appropriate rates [Schmidt et al., 2011]), then we could efficiently solve (3.28). This is precisely the case for the nonsmooth regularization functions mentioned in the previous section, whose proximity operators can be computed, in closed-form, as follows [Bach et al., 2011; Sra et al., 2011]:

- Lasso:

  $$[\text{prox}_{\lambda\|\cdot\|_1}(\mathbf{X})]_{ij} = \text{sgn}(\mathbf{X}_{ij})(|\mathbf{X}_{ij}| - \lambda)_+$$

  where $(\cdot)_+ = \max(\cdot, 0)$. This is the so-called componentwise soft-thresholding operator [Donoho & Johnstone, 1995].

- Group Lasso:

  $$[\text{prox}_{\lambda\|\cdot\|_{2,1}}(\mathbf{X})]_{i,:} = \frac{\mathbf{X}_{i,:}}{\|\mathbf{X}_{i,:}\|_2}(\|\mathbf{X}_{i,:}\|_2 - \lambda)_+$$

  where $\mathbf{X}_{i,:}$ denotes all the elements of the i-th row of the matrix $\mathbf{X}$.

- Sparse Group Lasso:

  $$[\text{prox}_{\lambda_1\|\cdot\|_1+\lambda\|\cdot\|_{2,1}}(\mathbf{X})]_{i,:} = \frac{\text{prox}_{\lambda_1\|\cdot\|_1}(\mathbf{V}_{i,:})}{\|\mathbf{V}_{i,:}\|_2}(\|\mathbf{V}_{i,:}\|_2 - \lambda)_+$$

- Trace norm:

  $$[\text{prox}_{\lambda\|\cdot\|_*}(\mathbf{X})] = \mathbf{U}(\mathbf{D} - \lambda)_+\mathbf{V}^\top$$

  where $\mathbf{X} = \mathbf{U}\mathbf{D}\mathbf{V}^\top$ is the SVD of the matrix $\mathbf{X}$.

Furthermore, when F has a Lipschitz continuous gradient, there are fast algorithms to solve (3.28). For instance, the Iterative Shrinkage Thresholding Algorithm (ISTA) has a convergence rate of $O(1/k)$, whereas the Fast Iterative Shrinkage-Thresholding Algorithm (FISTA), which is the algorithm used in this thesis, has a convergence rate of $O(1/k^2)$ [Beck & Teboulle, 2009].

3.4.2 Fast Iterative Shrinkage-Thresholding Algorithm

The Fast Iterative Shrinkage-Thresholding Algorithm (FISTA) [Beck & Teboulle, 2009] is an iterative method for solving the nonsmooth convex optimization problem (3.26). It requires that the smooth function $F(\mathbf{S})$ has a Lipschitz continuous gradient, meaning that:

$$\|\nabla F(\mathbf{S}_1) - \nabla F(\mathbf{S}_2)\|_2 \leq L\|\mathbf{S}_1 - \mathbf{S}_2\|_2 \tag{3.29}$$

where $\|\cdot\|_2$ is the Euclidean norm (equivalent to the Frobenius norm for matrices). This requirement is fulfilled by the smooth function involved in the regularized EEG inverse problem:

$$\begin{aligned} \|\nabla F(\mathbf{S}_1) - \nabla F(\mathbf{S}_2)\|_2^2 &= \|\mathbf{A}^\top(\mathbf{A}\mathbf{S}_1 - \mathbf{Y}) - \mathbf{A}^\top(\mathbf{A}\mathbf{S}_2 - \mathbf{Y})\|_2^2 \\ &\leq |||\mathbf{A}^\top\mathbf{A}|||_2^2\|\mathbf{S}_1 - \mathbf{S}_2\|_2^2 \\ \Rightarrow \|\nabla F(\mathbf{S}_1) - \nabla F(\mathbf{S}_2)\|_2 &\leq L\|\mathbf{S}_1 - \mathbf{S}_2\|_2 \end{aligned} \tag{3.30}$$

where $L = |||\mathbf{A}^\top\mathbf{A}|||_2$ is an upper bound on the Lipschitz constant of ∇F. FISTA works as follows [Beck & Teboulle, 2009]:

Algorithm 1 FISTA with constant stepsize

Require: L (an upper bound on the Lipschitz constant of ∇F)
initialization: $\mathbf{X} = \mathbf{S}_0 \in \mathbb{R}^{N\times T}$, $t_1 = 1$
repeat

$$\mathbf{S}_k = \text{prox}_{\frac{\lambda}{L}\Omega}(\mathbf{X}_k)$$
$$t_{k+1} = \frac{1 + \sqrt{1 + 4t_k^2}}{2}$$
$$\mathbf{X}_{k+1} = \mathbf{S}_k + \left(\frac{t_k - 1}{t_{k+1}}\right)(\mathbf{S}_k - \mathbf{S}_{k-1})$$

until stopping condition is met

Instead of working with a constant L, it is also possible to work with a local estimate of L, which can be obtained, for instance, using a backtracking search routine [Beck & Teboulle, 2009].

Chapter 4

Matrix factorization approach to solve the EEG inverse problem

In this chapter, we propose a new method to solve the EEG inverse problem, which takes into account the structured sparsity and low rank of the sources matrix $\mathbf{S}$. The method is based on the factorization of the matrix $\mathbf{S}$ as a product of a sparse coding matrix and a dense latent source matrix. The structured sparse-low-rank structure is enforced by minimizing a regularized functional that includes the $\ell_{2,1}$-norm of the coding matrix and the squared Frobenius norm of the latent source matrix. We develop an alternating optimization algorithm to solve the resulting nonsmooth-nonconvex minimization problem. We analyze the convergence of the optimization procedure, and we compare, under different synthetic scenarios, the performance of our method with respect to the Group Lasso and Trace Norm regularizers when they are applied directly to the target matrix.

4.1 Introduction

As we have seen in the previous chapter, the main task of the regularization function, used in the regularized EEG inverse problem, is to induce neurophysiological meaningful solutions, which take into account the structured sparsity of the sources matrix: during a particular cognitive task, only the sources related with the brain area involved in such a task will be active, that is, the sources matrix will have few nonzero rows. In this chapter, we propose a regularizer that takes into account, not only

the structured sparsity of the sources matrix $\mathbf{S}$, but also its low rank, capturing this way the linear relationship between the active sources and their corresponding neighbors. In order to do so, we propose a new method based on matrix factorization and regularization, with the aim of recovering the latent structure of the sources matrix. In the factorization, the first matrix, which acts as a coding matrix, is penalized using the $\ell_{2,1}$-norm, and the second one, which acts as a dense, full rank latent source matrix, is penalized using the squared Frobenius norm.

In our approach, the resulting optimization problem is nonsmooth and nonconvex. In order to handle the nonconvexity of the optimization problem, we use an iterative alternating minimization approach: minimizing over the coding matrix while maintaining fixed the latent source matrix and viceversa. Both optimization problems are convex: the first one can be solved using proximal splitting methods, while the second one can be solved directly in terms of a matrix inversion.

4.2 Problem formulation

In order to induce structured sparse-low-rank solutions, we propose to reformulate (3.2) using a matrix factorization approach, which involves expressing the matrix $\mathbf{S}$ as the product of two matrices, $\mathbf{S} = \mathbf{BC}$, obtaining the following nonlinear estimation model:

$$\mathbf{Y} = \mathbf{ABC} + \mathbf{E} \tag{4.1}$$

where $\mathbf{B}$ and $\mathbf{C}$ are penalized using the $\ell_{2,1}$-norm and the squared Frobenius norm, respectively. The resulting optimization problem can be stated as follows:

$$\begin{aligned}
\hat{\mathbf{B}}, \hat{\mathbf{C}} &= \underset{\mathbf{B},\mathbf{C}}{\operatorname{argmin}} \left\{ \frac{1}{2}\|\mathbf{A}(\mathbf{BC}) - \mathbf{Y}\|_{\mathrm{F}}^2 + \lambda \left(\sum_{i=1}^{N} \|\mathbf{B}(i,:)\|_2 + \frac{\rho}{2} \sum_{i=1}^{K} \|\mathbf{C}(i,:)\|_2^2 \right) \right\} \\
&= \underset{\mathbf{B},\mathbf{C}}{\operatorname{argmin}} \left\{ \frac{1}{2}\|\mathbf{A}(\mathbf{BC}) - \mathbf{Y}\|_{\mathrm{F}}^2 + \lambda \left(\|\mathbf{B}\|_{2,1} + \frac{\rho}{2}\|\mathbf{C}\|_{\mathrm{F}}^2 \right) \right\}
\end{aligned} \tag{4.2}$$

where $\lambda > 0$, $\rho > 0$, $\mathbf{B} \in \mathbb{R}^{N \times K}$, $\mathbf{C} \in \mathbb{R}^{K \times T}$, and $\mathbf{B}(i,:)$, $\mathbf{C}(i,:)$ denote the i-th row of $\mathbf{B}$ and $\mathbf{C}$, respectively. $K \ll \{N, T\}$, λ, and ρ are parameters of the model that must be adjusted.

In this formulation, which we denote as matrix factorization approach, the $\ell_{2,1}$-norm and the squared Frobenius norm induce structured sparsity and smoothness in the rows of $\mathbf{B}$ and $\mathbf{C}$, respectively, and therefore also in the rows of $\mathbf{S}$. Finally, the parameter K encloses the low rank of $\mathbf{S}$:

$$\begin{aligned} \operatorname{rank}(\mathbf{B}) &\leq \min\{N, K\} \Rightarrow \operatorname{rank}(\mathbf{B}) \leq K \\ \operatorname{rank}(\mathbf{C}) &\leq \min\{K, T\} \Rightarrow \operatorname{rank}(\mathbf{C}) \leq K \\ \operatorname{rank}(\mathbf{BC}) &\leq \min\{\operatorname{rank}(\mathbf{B}), \operatorname{rank}(\mathbf{C})\} \leq K \\ \Rightarrow \operatorname{rank}(\mathbf{S}) &\leq K \end{aligned}$$

Hence, the proposed regularization framework takes into account all the prior knowledge about the structure of the target matrix $\mathbf{S}$.

4.3 Optimization algorithm

4.3.1 Matrix factorization approach

In this section, we address the issue of implementing the learning method (4.2) numerically. We propose the following reparametrization of (4.2):

$$\begin{aligned} \mathbf{B} = \sqrt{\lambda\rho}\,\tilde{\mathbf{B}},\ \mathbf{C} = \frac{1}{\sqrt{\lambda\rho}}\,\tilde{\mathbf{C}} \quad &\Rightarrow \quad \mathbf{BC} = \left(\sqrt{\lambda\rho}\,\tilde{\mathbf{B}}\right)\left(\frac{1}{\sqrt{\lambda\rho}}\,\tilde{\mathbf{C}}\right) \\ &\Rightarrow \quad \mathbf{BC} = \tilde{\mathbf{B}}\tilde{\mathbf{C}} \end{aligned} \tag{4.3}$$

Using (4.3) in the objective function of (4.2), we get

$$\begin{aligned} \Rightarrow \quad & \frac{1}{2}\|\mathbf{A}(\tilde{\mathbf{B}}\tilde{\mathbf{C}}) - \mathbf{Y}\|_{\mathrm{F}}^2 + \lambda\left(\|\sqrt{\lambda\rho}\,\tilde{\mathbf{B}}\|_{2,1} + \frac{\rho}{2}\|\frac{1}{\sqrt{\lambda\rho}}\,\tilde{\mathbf{C}}\|_{\mathrm{F}}^2\right) \\ \Rightarrow \quad & \frac{1}{2}\|\mathbf{A}(\tilde{\mathbf{B}}\tilde{\mathbf{C}}) - \mathbf{Y}\|_{\mathrm{F}}^2 + \lambda\sqrt{\lambda\rho}\|\tilde{\mathbf{B}}\|_{2,1} + \frac{\lambda\rho}{2(\sqrt{\lambda\rho})^2}\|\tilde{\mathbf{C}}\|_{\mathrm{F}}^2 \\ \Rightarrow \quad & \frac{1}{2}\|\mathbf{A}(\tilde{\mathbf{B}}\tilde{\mathbf{C}}) - \mathbf{Y}\|_{\mathrm{F}}^2 + \tilde{\lambda}\|\tilde{\mathbf{B}}\|_{2,1} + \frac{1}{2}\|\tilde{\mathbf{C}}\|_{\mathrm{F}}^2 \end{aligned}$$

where $\tilde{\lambda} = \lambda\sqrt{\lambda\rho}$, and therefore, we get an optimization problem with only one regularization parameter:

$$\hat{\mathbf{B}}, \hat{\mathbf{C}} \quad = \quad \underset{\mathbf{B},\mathbf{C}}{\operatorname{argmin}} \left\{ \frac{1}{2}\|\mathbf{A}(\mathbf{BC}) - \mathbf{Y}\|_{\mathrm{F}}^2 + \lambda\|\mathbf{B}\|_{2,1} + \frac{1}{2}\|\mathbf{C}\|_{\mathrm{F}}^2 \,, \lambda > 0 \right\} \tag{4.4}$$

The optimization problem (4.4) is a simultaneous minimization over matrices $\mathbf{B}$ and $\mathbf{C}$. For a fixed $\mathbf{C}$, the minimum over $\mathbf{B}$ can be obtained using FISTA. On the other hand, for a fixed $\mathbf{B}$, the minimum over $\mathbf{C}$ can be solved directly in terms of a matrix inversion. These observations suggest an alternating minimization algorithm [Argyriou et al., 2008; Micchelli et al., 2013]:

Algorithm 2 Alternating minimization algorithm to solve the matrix factorization-regularization-based approach

Require: $\mathbf{Y} \in \mathbb{R}^{M \times T}$, $\mathbf{A} \in \mathbb{R}^{M \times N}$, $\mathbf{C_0} \in \mathbb{R}^{K \times T}$, K, λ

repeat

$$\mathbf{B_t} = \underset{\mathbf{B}}{\operatorname{argmin}} \left\{ \frac{1}{2} \|\mathbf{A}(\mathbf{B}\mathbf{C_{t-1}}) - \mathbf{Y}\|_{\mathrm{F}}^2 + \lambda \|\mathbf{B}\|_{2,1} + \frac{1}{2} \|\mathbf{C_{t-1}}\|_{\mathrm{F}}^2 \right\} \quad (4.5)$$

$$\mathbf{C_t} = \underset{\mathbf{C}}{\operatorname{argmin}} \left\{ \frac{1}{2} \|\mathbf{A}(\mathbf{B_t}\mathbf{C}) - \mathbf{Y}\|_{\mathrm{F}}^2 + \lambda \|\mathbf{B_t}\|_{2,1} + \frac{1}{2} \|\mathbf{C}\|_{\mathrm{F}}^2 \right\} \quad (4.6)$$

until stopping condition is met

In order to obtain the initialization matrix $\mathbf{C_0}$, we use an approach based on the singular value decomposition of $\mathbf{Y}$. Without loss of generality, let us work with (4.1) in the noiseless case:

$$\mathbf{Y} = \mathbf{ABC} \quad (4.7)$$

From (4.7), we can see that $\{\mathbf{Y_1}, \mathbf{Y_2}, \ldots, \mathbf{Y_M}\} \subset \text{Row Space}(\mathbf{C})$, where $\mathbf{Y_i}$ denotes the i-th row of $\mathbf{Y}$.

Now, let us obtain a rank-K approximation of $\mathbf{Y}$ by using a truncated SVD (truncated at the singular value σ_K):

$$\mathbf{Y} \approx \mathbf{U}_{M \times K} \mathbf{\Sigma}_{K \times K} \mathbf{V}_{K \times T}^{\top} \quad (4.8)$$

From the SVD theory [Horn & Johnson, 1990], we know that $\{\mathbf{Y_1}, \mathbf{Y_2}, \ldots, \mathbf{Y_M}\} \subset \text{Row Space}(\mathbf{V}^{\top})$; therefore, we can choose $\mathbf{C_0} = \mathbf{V}^{\top}$. Then, given $\mathbf{C_0}$, we can start iterating using (4.5) and (4.6).

Minimization over B (fixed C)

The minimization over $\mathbf{B}$ can be stated as follows:

$$\mathbf{B_t} = \underset{\mathbf{B}}{\operatorname{argmin}} \{F_{\mathbf{B}}(\mathbf{B}) + \lambda\|\mathbf{B}\|_{2,1}\ , \lambda > 0\} \tag{4.9}$$

where $F_{\mathbf{B}}(\mathbf{B}) = \frac{1}{2}\|\mathbf{A}(\mathbf{BC_{t-1}}) - \mathbf{Y}\|_{\mathrm{F}}^2 + \frac{1}{2}\|\mathbf{C_{t-1}}\|_{\mathrm{F}}^2$. This is a composite convex optimization problem involving the sum of a smooth function ($F_{\mathbf{B}}(\mathbf{B})$) and a nonsmooth function ($\lambda\|\mathbf{B}\|_{2,1}$). As we have seen in Section 3.4.1, this kind of problem can be efficiently handled using FISTA. In order to apply FISTA to solve (4.9), we first need to compute the following:

1. The gradient of the smooth function $F_{\mathbf{B}}(\mathbf{B})$

$$\nabla F_{\mathbf{B}}(\mathbf{B}) = \frac{\partial F_{\mathbf{B}}(\mathbf{B})}{\partial \mathbf{B}} = \mathbf{A}^\top(\mathbf{A}(\mathbf{B}\mathbf{C}_{t-1}) - \mathbf{Y})\mathbf{C}_{t-1}^\top$$

2. An upper bound on the Lipschitz constant (L) of $\nabla F_{\mathbf{B}}(\mathbf{B})$ (as was mentioned in Section 3.4.2, it can also be estimated using a backtracking search routine)

$$\begin{aligned}
\|\nabla F_{\mathbf{B}}(\mathbf{B}_1) - \nabla F_{\mathbf{B}}(\mathbf{B}_2)\|_2^2 &= \|\mathbf{A}^\top\mathbf{A}\mathbf{B}_1\mathbf{C}_{t-1}\mathbf{C}_{t-1}^\top - \mathbf{A}^\top\mathbf{A}\mathbf{B}_2\mathbf{C}_{t-1}\mathbf{C}_{t-1}^\top\|_2^2 \\
&= \|\mathbf{A}^\top\mathbf{A}(\mathbf{B}_1 - \mathbf{B}_2)\mathbf{C}_{t-1}\mathbf{C}_{t-1}^\top\|_2^2 \\
&= \sum_{j=1}^{K}\|(\mathbf{A}^\top\mathbf{A}(\mathbf{B}_1 - \mathbf{B}_2))(\mathbf{C}_{t-1}\mathbf{C}_{t-1}^\top)_j\|_2^2
\end{aligned}$$

where $(\mathbf{C}_{t-1}\mathbf{C}_{t-1}^\top)_j$ denotes the j-th column of the matrix $\mathbf{C}_{t-1}\mathbf{C}_{t-1}^\top$. Taking into account that $\|\mathbf{Q}x\|_2 \leq |||\mathbf{Q}|||_2\|x\|_2,\ \forall x \in \mathbb{R}^N, \forall \mathbf{Q} \in \mathbb{R}^{M\times N}$ [Horn & Johnson, 1990], we get

$$\begin{aligned}
\|\nabla F_{\mathbf{B}}(\mathbf{B}_1) - \nabla F_{\mathbf{B}}(\mathbf{B}_2)\|_2^2 &= \sum_{j=1}^{K}\|(\mathbf{A}^\top\mathbf{A}(\mathbf{B}_1 - \mathbf{B}_2))(\mathbf{C}_{t-1}\mathbf{C}_{t-1}^\top)_j\|_2^2 \\
&\leq \sum_{j=1}^{K}|||\mathbf{A}^\top\mathbf{A}(\mathbf{B}_1 - \mathbf{B}_2)|||_2^2\ \|(\mathbf{C}_{t-1}\mathbf{C}_{t-1}^\top)_j\|_2^2 \\
&\leq |||\mathbf{A}^\top\mathbf{A}(\mathbf{B}_1 - \mathbf{B}_2)|||_2^2 \sum_{j=1}^{K}\|(\mathbf{C}_{t-1}\mathbf{C}_{t-1}^\top)_j\|_2^2 \\
&\leq |||\mathbf{A}^\top\mathbf{A}(\mathbf{B}_1 - \mathbf{B}_2)|||_2^2\ \|\mathbf{C}_{t-1}\mathbf{C}_{t-1}^\top\|_2^2
\end{aligned} \tag{4.10}$$

From (4.10), taking into account that the spectral norm is submultiplicative ($|||\mathbf{PQ}|||_2 \leq |||\mathbf{P}|||_2 \; |||\mathbf{Q}|||_2, \; \forall \mathbf{P} \in \mathbb{R}^{M\times N}, \; \forall \mathbf{Q} \in \mathbb{R}^{N\times T}$), it follows that

$$\|\nabla F_{\mathbf{B}}(\mathbf{B}_1) - \nabla F_{\mathbf{B}}(\mathbf{B}_2)\|_2^2 \leq |||\mathbf{A}^\top\mathbf{A}|||_2^2 \; |||\mathbf{B}_1 - \mathbf{B}_2|||_2^2 \; \|\mathbf{C}_{t-1}\mathbf{C}_{t-1}^\top\|_2^2$$

and using the fact that $|||\mathbf{P}|||_2^2 \leq \|\mathbf{P}\|_2^2, \; \forall \mathbf{P} \in \mathbb{R}^{M\times N}$, we obtain

$$\begin{aligned} \|\nabla F_{\mathbf{B}}(\mathbf{B}_1) - \nabla F_{\mathbf{B}}(\mathbf{B}_2)\|_2^2 &\leq |||\mathbf{A}^\top\mathbf{A}|||_2^2 \; \|\mathbf{B}_1 - \mathbf{B}_2\|_2^2 \; \|\mathbf{C}_{t-1}\mathbf{C}_{t-1}^\top\|_2^2 \\ &\leq L \; \|\mathbf{B}_1 - \mathbf{B}_2\|_2^2 \end{aligned} \tag{4.11}$$

where $L = |||\mathbf{A}^\top\mathbf{A}|||_2 \; \|\mathbf{C}_{t-1}\mathbf{C}_{t-1}^\top\|_2$.

3. Finally, we need to compute the proximal operator associated to the nonsmooth function $\lambda\|\cdot\|_{2,1}$, whose closed-form expression was given in Section 3.4.1.

Minimization over C (fixed B)

The minimization over $\mathbf{C}$ can be stated as follows:

$$\mathbf{C_t} = \underset{\mathbf{C}}{\operatorname{argmin}} \{F_{\mathbf{C}}(\mathbf{C})\} \tag{4.12}$$

where $F_{\mathbf{C}}(\mathbf{C}) = \frac{1}{2}\|\mathbf{A}(\mathbf{B_t}\mathbf{C}) - \mathbf{Y}\|_{\mathrm{F}}^2 + \lambda\|\mathbf{B_t}\|_{2,1} + \frac{1}{2}\|\mathbf{C}\|_{\mathrm{F}}^2$ is a smooth function of $\mathbf{C}$. In what follows, we show how the minimum over $\mathbf{C}$ can be solved directly in terms of a matrix inversion:

$$\begin{aligned} \nabla F_{\mathbf{C}}(\mathbf{C}) = 0 \;\; &\Rightarrow \;\; \mathbf{B_t}^\top\mathbf{A}^\top(\mathbf{A}(\mathbf{B_t}\mathbf{C}_t) - \mathbf{Y}) + \mathbf{C_t} = 0 \\ &\Rightarrow \;\; \mathbf{C_t} = \left[\mathbf{B_t}^\top\mathbf{A}^\top\mathbf{A}\mathbf{B_t} + \mathbf{I}_K\right]^{-1}\mathbf{B_t}^\top\mathbf{A}^\top\mathbf{Y} \end{aligned} \tag{4.13}$$

The matrix $\left[\mathbf{B_t}^\top\mathbf{A}^\top\mathbf{A}\mathbf{B_t} + \mathbf{I}_K\right] \in \mathbb{R}^{K\times K}$, and K is supposed to be small; therefore, calculating its corresponding inverse matrix is quite cheap.

4.4 Convergence analysis

We are going to analyze the convergence behavior of Algorithm 2 by using the global convergence theory of iterative algorithms developed by Zangwill [Zangwill, 1969]. Note that in this theory, the term 'global convergence' does not imply convergence to a global optimum for all initial points. The property of global convergence expresses, in a sense, the certainty that the algorithm converges to the solution set. Formally, an iterative algorithm ξ, on the set X, is said to be globally convergent provided, for any starting point $x_0 \in X$, the sequence $\{x_n\}$ generated by ξ has a limit point [Sriperumbudur & Lanckriet, 2009].

Before we state the Zangwill's global convergence theorem, we need a formal definition of iterative algorithm, as well as the definition of a set-valued mapping (point-to-set mapping) [Zangwill, 1969]:

Definition 4.4.1. *Set-valued mapping. Given two sets, X and Y, a set-valued mapping defined on X, with range in the power set of Y, $\mathcal{P}(Y)$, is a map, Φ, which assigns to each $x \in X$ a subset $\Phi(x) \in \mathcal{P}(Y)$,*

$$\Phi : X \to \mathcal{P}(Y)$$

Definition 4.4.2. *Iterative algorithm. Let X be a set and $x_0 \in X$ a given point. Then, an iterative algorithm ξ, with initial point x_0, is a set-valued mapping*

$$\xi : X \to \mathcal{P}(X)$$

which generates a sequence $\{x_n\}_{n=1}^{\infty}$ via the rule $x_{n+1} \in \xi(x_n)$, $n = 0, 1, \ldots$

Theorem 4.4.1. *[Zangwill, 1969] Zangwill's global convergence theorem. Let the set-valued mapping $M_x(x) : X \to \mathcal{P}(X)$ determine an algorithm that given a point x_0 generates a sequence $\{x_n\}_{n=0}^{\infty}$ through the iteration $x_{n+1} \in M_x(x_n)$. Also, let a solution set Γ be given. Suppose*

1. *All point x_n are in a compact set $S \subset X$.*
2. *There is a continuous function $\alpha : X \to \mathbb{R}$ such that*
   (a) *if $x \notin \Gamma$, then $\alpha(x') < \alpha(x)$ $\forall x' \in M_x(x)$.*
   (b) *if $x \in \Gamma$, then $\alpha(x') \leq \alpha(x)$ $\forall x' \in M_x(x)$.*
3. *The map $M_x(x)$ is closed at x if $x \notin \Gamma$.*

Then, the limit of any convergent subsequence of $\{x_n\}_{n=0}^{\infty}$ *is in* Γ*. That is, accumulation points* x^* *of the sequence* x_n *lie in* Γ*. Furthermore,* $\alpha(x_n)$ *converges to* α^**, and* $\alpha(x^*)=\alpha^*$ *for all accumulation points* x^**.*

In order to use Theorem 4.4.1 to analyze the convergence properties of Algorithm 2, we will need the following definitions and theorems:

Definition 4.4.3. *Compact set. A set* X *is said to be compact if any sequence (or subsequence) contains a convergent subsequence whose limit is in* X*. More explicitly, given a subsequence* $\{x_n\}_{n\in\widehat{N}}$ *in* X*, there exists a* $\widehat{N_1} \subset \widehat{N}$ *such that*

$$x_n \to x_\infty, \quad n \in \widehat{N_1}$$

with $x_\infty \in X$ *(we write convergence of subsequences as* $x_n \to x_\infty$*, which is equivalent to* $\lim_{n\to\infty} x_n = x_\infty$*).*

Definition 4.4.4. *Composite map. Let* $\Pi_A : X \to Y$ *and* $\Pi_B : Y \to Z$ *be two set-valued mappings. The composite map* $\Pi_C = \Pi_B \circ \Pi_A$ *which takes points* $x \in X$ *to sets* $\Pi_C(x) \subset Z$ *is defined by*

$$\Pi_C(x) := \bigcup_{y\in\Pi_A(x)} \Pi_B(y)$$

Definition 4.4.5. *Closed map. A set-valued mapping* $\Pi : X \to \mathcal{P}(Y)$ *is closed at* $x_0 \in X$ *provided*

1. $x_n \to x_0$ *as* $n \to \infty$, $x_n \in X$
2. $y_n \to y_0$ *as* $n \to \infty$, y_n, $y_0 \in Y$
3. $y_n \in \Pi(x_n)$

implies $y_0 \in \Pi(x_0)$*. The map* Π *is called closed on* $S \subset X$ *provided it is closed at each* $x \in S$*.*

Theorem 4.4.2. *Composition of closed maps. Let* $\Pi_A : X \to Y$ *and* $\Pi_B : Y \to Z$ *be two set-valued mappings. Suppose*

1. Π_A *is closed at* x_0
2. Π_B *is closed on* $\Pi_A(x_0)$
3. *if* $x_n \to x_0$ *and* $y_n \in \Pi_A(x_n)$*, then there exists* $y_0 \in Y$*, such that for some sequence* $\{y_{n_j}\}$, $y_{n_j} \to y_0$ *as* $j \to \infty$*.*

Then, the composite map $\Pi_C = \Pi_B \circ \Pi_A$ is closed at x_0.

Lemma 4.4.1. *[Gunawardana & Byrne, 2005] Given a real-valued function defined on $X \times Y$, define the set-valued mapping $\Psi : X \to \mathcal{P}(Y)$ by*

$$\Psi(x) = \underset{y \in Y}{\operatorname{argmin}}\, h(x, y)$$

then, Ψ is closed at x if $\Psi(x)$ is nonempty.

Theorem 4.4.3. *Weierstrass theorem. If f is a real continuous function on a compact set $S \subset \mathbb{R}^n$, then the problem*

$$\underset{x \in \mathbb{R}^n}{\operatorname{argmin}} \{f(x),\ x \in S\}$$

has an optimal solution $x^* \in S$.

Now that we know the main building blocks of the global convergence theory of iterative algorithms, we are in a position to state the convergence theorem related to Algorithm 2. This convergence theorem is a direct application of Theorem 4.4.1:

Theorem 4.4.4. *Let Φ denotes the iterative Algorithm 2, and suppose that given $\mathbf{Y} \in \mathbb{R}^{M \times T}$, $\mathbf{A} \in \mathbb{R}^{M \times N}$, $\mathbf{B_0} \in \mathbb{R}^{N \times K}$, $\mathbf{C_0} \in \mathbb{R}^{K \times T}$, K, and λ, the sequence $\{\mathbf{B_t}, \mathbf{C_t}\}_{t=1}^{\infty}$ is generated and satisfies $\{\mathbf{B_{t+1}}, \mathbf{C_{t+1}}\} \in \Phi(\mathbf{B_t}, \mathbf{C_t})$. Also, let Ω_B and Ω_C denote the solution sets of (4.5) and (4.6), respectively:*

$$\Omega_B = \left\{ \mathbf{B} \in \mathbb{R}^{N \times K} \;\middle|\; 0 \in \partial \left(\frac{1}{2} ||\mathbf{A}(\mathbf{B}\mathbf{C_{t-1}}) - \mathbf{Y}||_{\mathrm{F}}^2 + \lambda ||\mathbf{B}||_{2,1} + \frac{1}{2} ||\mathbf{C_{t-1}}||_{\mathrm{F}}^2 \right) \right\}$$

$$\Omega_C = \left\{ \mathbf{C} \in \mathbb{R}^{K \times T} \;\middle|\; \nabla \left(\frac{1}{2} ||\mathbf{A}(\mathbf{B_t}\mathbf{C}) - \mathbf{Y}||_{\mathrm{F}}^2 + \lambda ||\mathbf{B_t}||_{2,1} + \frac{1}{2} ||\mathbf{C}||_{\mathrm{F}}^2 \right) = 0 \right\}$$

Then, the limit of any convergent subsequence of $\{\mathbf{B_t}, \mathbf{C_t}\}_{t=1}^{\infty}$ is in Ω_B and Ω_C.

Proof. The iterative algorithm Φ can be decomposed into two well-defined iterative algorithms Φ_B and Φ_C:

$$\Phi_B(\mathbf{C_{t-1}}) = \mathbf{B_t} = \underset{\mathbf{B}}{\operatorname{argmin}} \left\{ \frac{1}{2} ||\mathbf{A}(\mathbf{B}\mathbf{C_{t-1}}) - \mathbf{Y}||_{\mathrm{F}}^2 + \lambda ||\mathbf{B}||_{2,1} + \frac{1}{2} ||\mathbf{C_{t-1}}||_{\mathrm{F}}^2 \right\} \tag{4.14}$$

$$\Phi_C(\mathbf{B_t}) = \mathbf{C_t} = \underset{\mathbf{C}}{\operatorname{argmin}} \left\{ \frac{1}{2} ||\mathbf{A}(\mathbf{B_t}\mathbf{C}) - \mathbf{Y}||_{\mathrm{F}}^2 + \lambda ||\mathbf{B_t}||_{2,1} + \frac{1}{2} ||\mathbf{C}||_{\mathrm{F}}^2 \right\} \tag{4.15}$$

As we can see from (4.14) and (4.15), at iteration t, the result of Φ_B becomes the input of Φ_C, and at iteration $t+1$, the result of Φ_C becomes the input of Φ_B; therefore, we can express Φ as the composition of Φ_C and Φ_B, that is, $\Phi(\mathbf{C_{t-1}}) = \Phi_C(\Phi_B(\mathbf{C_{t-1}}))$:

$$\Phi_C(\Phi_B(\mathbf{C_{t-1}})) = \mathbf{C_t} = \underset{\mathbf{C}}{\operatorname{argmin}} \left\{ \frac{1}{2}||\mathbf{A}(\mathbf{B_t C}) - \mathbf{Y}||_{\mathrm{F}}^2 + \lambda||\mathbf{B_t}||_{2,1} + \frac{1}{2}||\mathbf{C}||_{\mathrm{F}}^2 \right\} \tag{4.16}$$

$$\text{subject to} \quad \mathbf{B_t} = \underset{\mathbf{B}}{\operatorname{argmin}} \left\{ \frac{1}{2}||\mathbf{A}(\mathbf{B C_{t-1}}) - \mathbf{Y}||_{\mathrm{F}}^2 + \lambda||\mathbf{B}||_{2,1} + \frac{1}{2}||\mathbf{C_{t-1}}||_{\mathrm{F}}^2 \right\}$$

Let Γ be the solution set of Φ:

$$\Gamma = \left\{ \mathbf{C} \in \mathbb{R}^{K \times T} \;\middle|\; \frac{\partial Z(\mathbf{C}, t)}{\partial \mathbf{C}} = 0 \right\}$$

where $Z(\mathbf{C}, t) = \frac{1}{2}||\mathbf{A}(\mathbf{B_t C}) - \mathbf{Y}||_{\mathrm{F}}^2 + \lambda||\mathbf{B_t}||_{2,1} + \frac{1}{2}||\mathbf{C}||_{\mathrm{F}}^2$.

To prove this theorem by using Zangwill's global convergence theorem, we need to prove that all its corresponding assumptions are fulfilled. In order to prove assumption 1, let us analyze the sequences $\{\mathbf{B_t}\}_{t=1}^{\infty}$ and $\{\mathbf{C_t}\}_{t=1}^{\infty}$. The sequence $\{\mathbf{B_t}\}_{t=1}^{\infty}$ is generated using FISTA, which is a convergent algorithm ($\mathbf{B_t} \to \mathbf{B}_\infty$) that guarantees that $\mathbf{B_t} \in \Omega_B$ [Beck & Teboulle, 2009; Combettes & Wajs, 2005]. Hence, using Definition 4.4.3, we can see that the sequence $\{\mathbf{B_t}\}_{t=1}^{\infty}$ generated by (4.14) lies in a compact set. On the other hand, the sequence $\{\mathbf{C_t}\}_{t=1}^{\infty}$ is generated by (4.13), which guarantees that $\mathbf{C_t} \in \Omega_C$. This sequence always converges to a point inside Ω_C, which implies that Ω_C also lies in a compact set. This concludes the proof of assumption 1.

To prove assumption 2, let us use $Z(\mathbf{C}, t)$ as the function $\alpha(\cdot)$; thus, in order to verify the fulfillment of assumption 2, we need to prove that

(a) if $\mathbf{C_t} \notin \Gamma$, then $Z(\mathbf{C_{t+1}}, t+1) < Z(\mathbf{C_t}, t)\ \forall \mathbf{C_{t+1}} \in \Phi(\mathbf{C_t})$

(b) if $\mathbf{C_t} \in \Gamma$, then $Z(\mathbf{C_{t+1}}, t+1) \leq Z(\mathbf{C_t}, t)\ \forall \mathbf{C_{t+1}} \in \Phi(\mathbf{C_t})$

From (4.16), we can see that the sequence $\{\mathbf{C_t}\}_{t=1}^{\infty}$ will always lie in Γ (because $\mathbf{C_t}$ is generated by (4.13)); therefore, we only need to prove (b).

Let $\mathbf{C_{t+1}}$ be the solution of (4.16) at iteration $t+1$; this implies

$$\begin{aligned}\frac{1}{2}||\mathbf{A}(\mathbf{B_{t+1}C_{t+1}})-\mathbf{Y}||_{\mathrm{F}}^2+\lambda||\mathbf{B_{t+1}}||_{2,1}+\frac{1}{2}||\mathbf{C_{t+1}}||_{\mathrm{F}}^2 \\ \leq \frac{1}{2}||\mathbf{A}(\mathbf{B_{t+1}C})-\mathbf{Y}||_{\mathrm{F}}^2+\lambda||\mathbf{B_{t+1}}||_{2,1}+\frac{1}{2}||\mathbf{C}||_{\mathrm{F}}^2,\ \forall \mathbf{C}\in\mathbb{R}^{K\times T} \\ \leq \frac{1}{2}||\mathbf{A}(\mathbf{B_{t+1}C_t})-\mathbf{Y}||_{\mathrm{F}}^2+\lambda||\mathbf{B_{t+1}}||_{2,1}+\frac{1}{2}||\mathbf{C_t}||_{\mathrm{F}}^2\end{aligned} \tag{4.17}$$

On the other hand, if $\mathbf{B_{t+1}}$ is the solution of (4.14) at iteration $t+1$, this implies

$$\begin{aligned}\frac{1}{2}||\mathbf{A}(\mathbf{B_{t+1}C_t})-\mathbf{Y}||_{\mathrm{F}}^2+\lambda||\mathbf{B_{t+1}}||_{2,1}+\frac{1}{2}||\mathbf{C_t}||_{\mathrm{F}}^2 \\ \leq \frac{1}{2}||\mathbf{A}(\mathbf{BC_t})-\mathbf{Y}||_{\mathrm{F}}^2+\lambda||\mathbf{B}||_{2,1}+\frac{1}{2}||\mathbf{C_t}||_{\mathrm{F}}^2,\ \forall \mathbf{B}\in\mathbb{R}^{N\times K} \\ \leq \frac{1}{2}||\mathbf{A}(\mathbf{B_tC_t})-\mathbf{Y}||_{\mathrm{F}}^2+\lambda||\mathbf{B_t}||_{2,1}+\frac{1}{2}||\mathbf{C_t}||_{\mathrm{F}}^2\end{aligned} \tag{4.18}$$

and from (4.17) and (4.18), we can prove assumption 2(b):

$$\begin{aligned}\frac{1}{2}||\mathbf{A}(\mathbf{B_{t+1}C_{t+1}})-\mathbf{Y}||_{\mathrm{F}}^2+\lambda||\mathbf{B_{t+1}}||_{2,1}+\frac{1}{2}||\mathbf{C_{t+1}}||_{\mathrm{F}}^2 \\ \leq \frac{1}{2}||\mathbf{A}(\mathbf{B_tC_t})-\mathbf{Y}||_{\mathrm{F}}^2+\lambda||\mathbf{B_t}||_{2,1}+\frac{1}{2}||\mathbf{C_t}||_{\mathrm{F}}^2 \\ Z(\mathbf{C_{t+1}},t+1)\leq Z(\mathbf{C_t},t)\end{aligned}$$

In order to prove assumption 3, we need to prove that Φ is closed at $\mathbf{C}$ if $\mathbf{C}\notin\Gamma$. To do so, we are going to use Theorem 4.4.2; therefore, we need to prove that Φ_B and Φ_C are both closed maps: from (4.14) and (4.15), we can see that their corresponding objective functions are both continuous $\forall\mathbf{B}\in\mathbb{R}^{N\times K}$ and $\forall\mathbf{C}\in\ \mathbb{R}^{K\times T}$, respectively; hence, by using Weierstrass Theorem and Lemma 4.4.1, we can conclude that Φ_B and Φ_C are both closed maps for any $\mathbf{C_{t-1}}$ and $\mathbf{B_t}$, respectively, and by using Theorem 4.4.2, we can conclude that Φ is closed on any $\mathbf{C_{t-1}}$.

Finally, from all the previous proofs and Zangwill's global convergence theorem, it follows that the limit of any convergent subsequence of $\{\mathbf{B_t},\mathbf{C_t}\}_{t=1}^{\infty}$ is in Ω_B and Ω_C.

□

4.5 Numerical experiments

In this section, we evaluate the performance of the matrix factorization approach and compare it with the Group Lasso regularizer:

$$\hat{\mathbf{S}} = \underset{\mathbf{S}}{\operatorname{argmin}} \left\{ \frac{1}{2}\|\mathbf{AS}-\mathbf{Y}\|_{\mathrm{F}}^2 + \lambda \sum_{i=1}^{N} \|\mathbf{S}(i,:)\|_2 \, , \lambda > 0 \right\} \quad (4.19)$$

and the Trace Norm regularizer:

$$\hat{\mathbf{S}} = \underset{\mathbf{S}}{\operatorname{argmin}} \left\{ \frac{1}{2}\|\mathbf{AS}-\mathbf{Y}\|_{\mathrm{F}}^2 + \lambda \sum_{i=1}^{q} \sigma_i(\mathbf{S}) \, , \lambda > 0 \right\} \quad (4.20)$$

where $q = \min\{\mathrm{N}, \mathrm{T}\}$ and $\sigma_i(\mathbf{S})$ denotes the i-th singular value of $\mathbf{S}$. Both problems (4.19) and (4.20) were solved using the FISTA implementation of the SPArse Modeling Software (SPAMS) [Jenatton et al., 2010; Mairal et al., 2010].

In order to have a reproducible comparison of the different regularization approaches, we generated two synthetic scenarios:

- $M = 128$ EEG electrodes, $T = 161$ time instants, $N = 413$ current sources within the brain, but only 12 of them are active: 4 main active sources with their corresponding 2 nearest neighbor sources also active. The other 401 sources are not active (zero electrical activity). Therefore, in this scenario, the synthetic matrix $\mathbf{S}$ is a structured sparse matrix with only 12 nonzero rows (the rows associated to the active sources).

- $M = 128$ EEG electrodes, $T = 161$ time instants, $N = 2052$ current sources within the brain, but only 40 of them are active: 4 main active sources with their corresponding 9 nearest neighbor sources also active. The other 2012 sources are not active (zero electrical activity). Therefore, in this scenario, the synthetic matrix $\mathbf{S}$ is a structured sparse matrix with only 40 nonzero rows (the rows associated to the active sources).

In both scenarios, the simulated electrical activity (simulated waveforms) associated to the four Main Active Sources (MAS) was obtained from a face perception-evoked potential study [Friston et al., 2008; Henson et al., 2003]. To obtain the simulated electrical activity associated to each one of the active neighbor sources, we simply set it as a scaled version of the

electrical activity of its corresponding nearest MAS (with a scaled factor equal to 0.5). Hence, there is a linear relation between the four MAS and their corresponding nearest neighbor sources; therefore, in both scenarios, the rank of the synthetic matrix $\mathbf{S}$ is equal to 4.

As forward model, we used a three-shell spherical head model. To obtain the values of each one of the components of the lead-field matrix $\mathbf{A}$, we solved the EEG forward problem using the SPM software [Litvak et al., 2011].

Finally, the simulated EEG signals were generated according to (3.2), where $\mathbf{E}$ is a Gaussian noise $G(0, \sigma^2\mathbf{I})$ whose variance was set to satisfy a Signal to Noise Ratio (SNR) $= 20\log_{10}\left(\frac{||\mathbf{AS}||_{\mathrm{F}}}{||\mathbf{E}||_{\mathrm{F}}}\right) = 10$ dB. Summarizing, our synthetic problems can be stated as follows: Given matrices $\mathbf{Y} \in \mathbb{R}^{128\times 161}$ and $\mathbf{A} \in \mathbb{R}^{128\times N}$, recover the synthetic sources matrix $\mathbf{S} \in \mathbb{R}^{N\times 161}$. According to this, in both scenarios, we want to estimate a sources matrix which is structured sparse and low rank, with its rank equal to the number of MAS simulated. The activity of the four MAS, the synthetic EEG measurements as well as the sparsity pattern of the synthetic sources matrix are shown in Figures 4.1 and 4.2 (Ground Truth).

We have used cross-validation to select the regularization parameter λ associated to the Group Lasso and Trace Norm regularizers, as well as the parameters λ and K in the case of the Matrix Factorization approach ($K \in [1, 2, 3, \ldots, 10]$, $\lambda \in [10^{-3}, 10^{-2}, 10^{-1}, \ldots, 10^{3}]$): the rows of $\mathbf{Y}$ are randomly partitioned into three groups of approximately equal size. Each union of two groups forms a train set (TrS), while the remaining group forms a test set (TS). This procedure is carried out three times, each time selecting a different test group. Inverse reconstructions are carried out based on the training sets, obtaining different regression matrices $\hat{\mathbf{S}}_i$. We then evaluate the root mean square error (RMSE) using the test sets and the regression matrices $\hat{\mathbf{S}}_i$:

$$\mathrm{RMSE} : \frac{1}{3}\sum_{i=1}^{3}\left(\frac{1}{\sqrt{M_{\mathrm{TS}_i}\times T}}\|\mathbf{A}_{\mathrm{TS}_i}\hat{\mathbf{S}}_i - \mathbf{Y}_{\mathrm{TS}_i}\|_{\mathrm{F}}\right)$$

where $\mathbf{Y}_{\mathrm{TS}_i} \in \mathbb{R}^{M_{\mathrm{TS}_i}\times T}$, and $\mathbf{A}_{\mathrm{TS}_i} \in \mathbb{R}^{M_{\mathrm{TS}_i}\times N}$ (TS_i denotes the index set of the rows that belongs to the i-th test set). Once the estimated matrix $\hat{\mathbf{S}}$ has been found, we apply a threshold to remove spurious sources with almost zero activity. We have set this threshold equal to the 1% of the mean energy of all the sources.

4.5.1 Performance evaluation

In order to evaluate the performance of the regularizers, we compare the waveform and localization of the four MAS present in the synthetic sources matrix against the four MAS estimated by each one of the regularizers. We also compare the sparsity pattern of the estimated sources matrix $\hat{\mathbf{S}}$ against the sparsity pattern of the synthetic sources matrix $\mathbf{S}$, as well as the synthetic and predicted EEG measurements.

As we can see from Figures 4.1 and 4.2, the Group Lasso and Trace Norm regularizers do not reveal the correct number of linear independent sources, while the Matrix Factorization does: it finds out four linear independent sources in both scenarios. To select such four linear independent MAS, we find a basis for the Column Space($\tilde{\mathbf{S}}^\top$) (using a QR factorization), where $\tilde{\mathbf{S}}$ is a matrix whose rows are a sorted version of the rows of $\mathbf{S}$ (sorted in a descending order of their corresponding energy value). To get the four linear independent MAS estimated by the Group Lasso and Trace Norm regularizers, we followed the same procedure described before and retained the first four components of the basis of the Column Space($\tilde{\mathbf{S}}^\top$).

According to Figures 4.1 and 4.2, the Matrix Factorization approach is able to estimate a sources matrix with the correct rank and whose sparsity pattern follows closely the sparsity pattern of the true sources matrix, that is, both matrices have a similar structure, which implies that the proposed approach is able to induce the desired solution: a row-structured sparse matrix, whose nonzero rows encode the linear relationship between the active sources and their corresponding nearest neighbor sources. Using the estimated sources matrix $\hat{\mathbf{S}}$, the Matrix Factorization approach is also able to predict a smooth version of the noisy EEG, and the waveforms of the estimated MAS follow closely the waveforms of the true MAS.

As we can see from Figures 4.1 and 4.2, Group Lasso is able to estimate a sources matrix with a similar row-sparsity pattern to the true sources matrix, but it does not take into account the linear relationship between the nonzero rows, which can be seen from the rank of the estimated sources matrix. The waveforms of the estimated MAS are very similar to the true MAS, but they are not so smooth as the ones estimated by the Matrix Factorization approach.

As we can see from Figures 4.1 and 4.2, the Trace Norm regularizer takes into account the linear relationship of the active sources by inducing solutions which are low rank, but, on the other hand, it does not take into account the structured sparsity pattern of the sources matrix. All of this implies that the Trace Norm tends to induce low rank dense solutions, which

are not biologically plausible.

According to Figures 4.3(a) and 4.3(b), the position of the MAS, obtained from the matrices estimated using the Matrix Factorization approach, the Group Lasso, and Trace Norm regularizers, respectively, follows closely the position of the true MAS. Nevertheless, it is worth to highlight that before selecting the MAS, we first need an accurate estimation of their number, and the Group Lasso and Trace Norm regularizers were not able to get a precise estimate of it, only the Matrix Factorization were able to.

From these results, we can see that the proposed Matrix Factorization approach outperforms both the Group Lasso and Trace Norm regularizers. The main reason for this is because it combines their two main features: it combines the structured sparsity (from Group Lasso) and the low rank (from Trace Norm) into one unified framework, which implies that it is able to induce structured sparse-low-rank solutions which are biologically plausible: few active sources, with linear relationships between them.

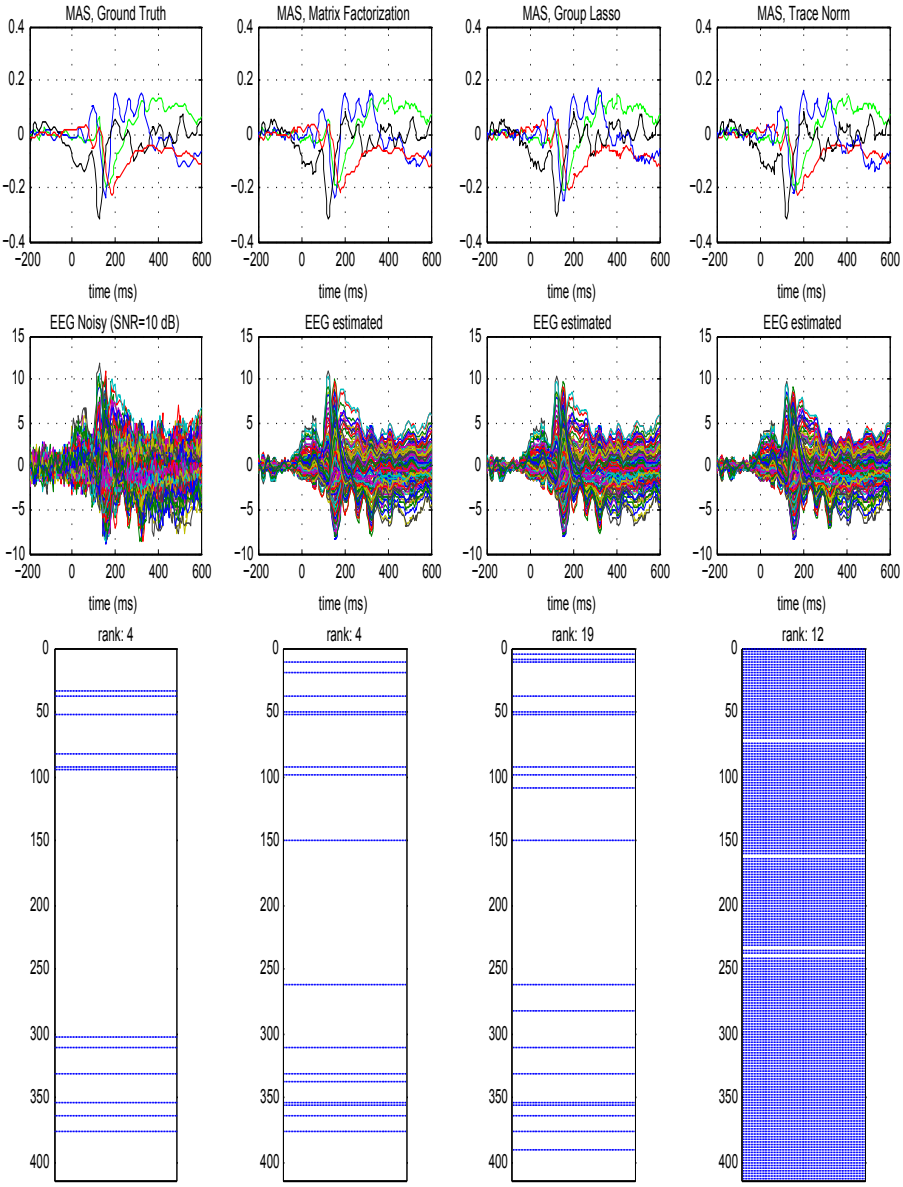


Figure 4.1: Simulation results: waveforms of the MAS, EEG estimated and sparsity pattern of the estimated sources matrix $\hat{\mathbf{S}}$. Experiment setup: 413 sources, 128 EEG electrodes, 161 time instants, 4 main active sources with their corresponding 2 nearest neighbor sources also active.

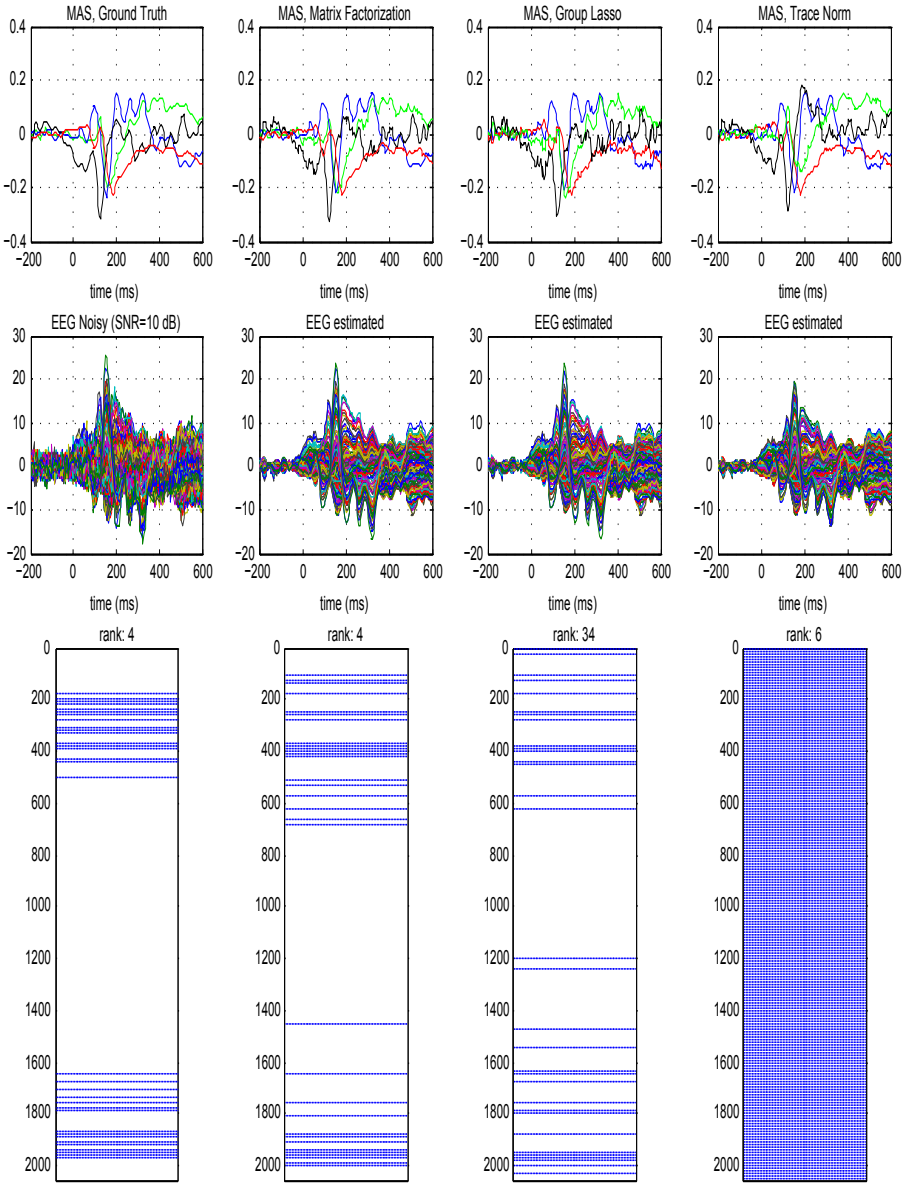
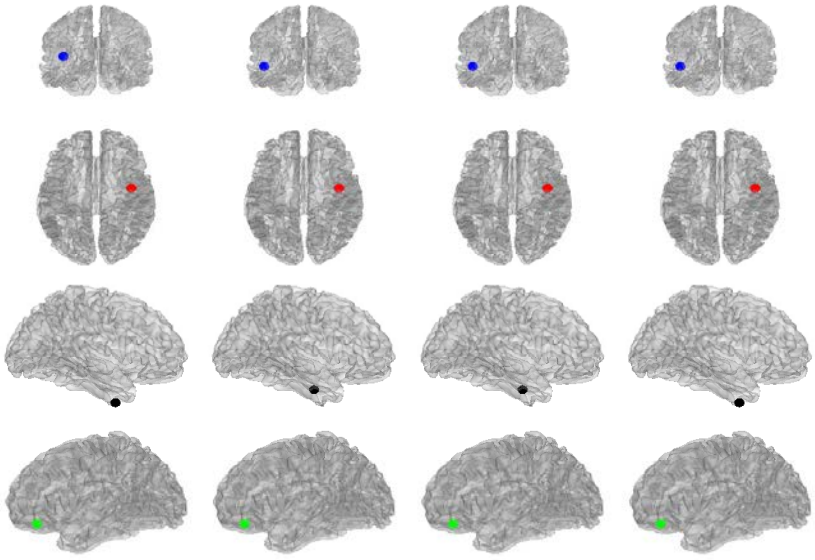
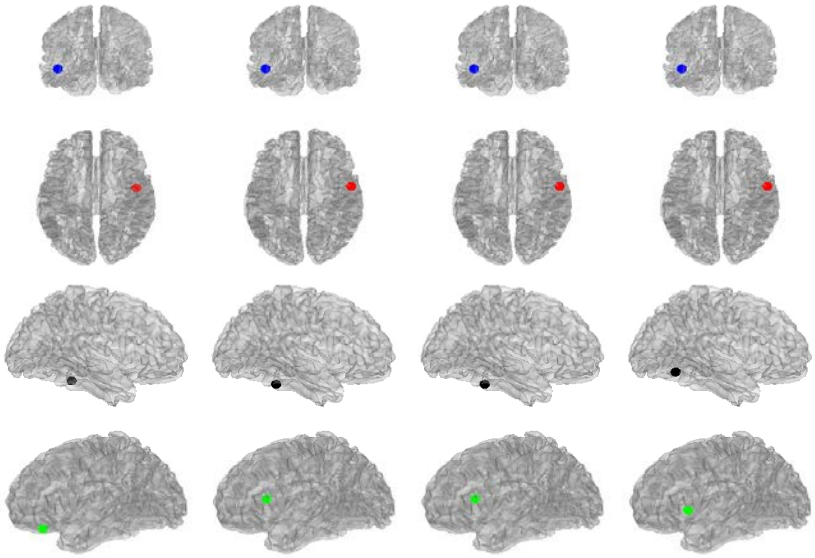


Figure 4.2: Simulation results: waveforms of the MAS, EEG estimated and sparsity pattern of the estimated sources matrix $\hat{\mathbf{S}}$. Experiment setup: 2052 sources, 128 EEG electrodes, 161 time instants, 4 main active sources with their corresponding 9 nearest neighbor sources also active.



(a) N=413 sources.



(b) N=2052 sources.

Figure 4.3: Localization of the MAS. From left to right: Ground Truth, Matrix Factorization, Group Lasso and Trace Norm.

Chapter 5

Mobile brain scanners

Mobile brain scanning is a new neuroinformatics paradigm that combines a wireless EEG neuroheadset, in charge of acquiring and transmitting the electrical potentials measured on the scalp, with a mobile device, in charge of receiving and processing these data to generate 3D cortical activation maps, which show us the brain regions that are currently active. To generate such activation maps, the mobile brain scanner needs to solve, as fast as possible, the EEG inverse problem. Therefore, the EEG inverse solver is the most important component of the system: it determines how fast is the scanning process and how relieble are the active sources that it estimates. In this chapter, we describe each one of the hardware and software components of a mobile brain scanner and how they work together to generate the cortical activation maps. We also mention the components that we have selected to develop a working prototype of a mobile brain scanner, whose implementation details will be explained in the following chapter.

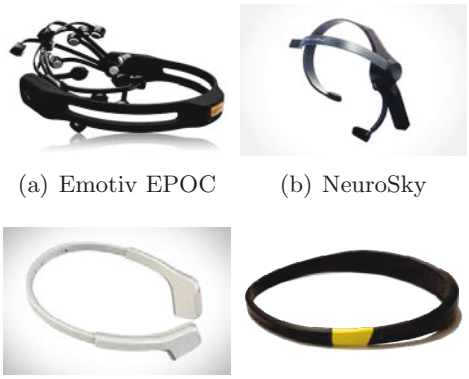
5.1 Mobile brain scanner: hardware components

The hardware components of a mobile brain scanner consist in a wireless EEG neuroheadset and one mobile device (smartphone or tablet). Unlike fMRI, this mobile system can be used to study the brain on a wide variety of scenarios, in which the person being scanned can move and act under naturalistic conditions. Besides, it has direct access to the brain electrical activity, which is useful to study the dynamics of the brain processes on the millisecond timescale on which neurons operate.

Figure 5.1: Main components of a mobile brain scanning system.

5.1.1 Wireless EEG neuroheadset

The wireless EEG neuroheadset is in charge on acquiring and transmitting the electrical potentials measured on the scalp. Recently, several low-cost wireless EEG neuroheadsets have been made available, including the Emotiv EPOC neuroheadset[1], NeuroSky[2], Muse[3] and Melon[4] (see Figure 5.2). These EEG platforms support applications ranging from emotion recognition, Brain Computer Interface (BCI), game control, stress reduction, cognitive training, to sleep monitoring.



(a) Emotiv EPOC (b) NeuroSky

(c) Muse (d) Melon

Figure 5.2: Wireless EEG neuroheadsets.

Throughout this thesis we use the Emotiv EPOC neuroheadset. This

[1]http://www.emotiv.com
[2]http://neurosky.com
[3]http://www.choosemuse.com
[4]http://www.thinkmelon.com

EEG neuroheadset includes 14 EEG sensors, see Figure 5.3, aligned with the 10-20 system (AF3, F7, F3, FC5, T7, P7, O1, O2, P8, T8, FC6, F4, F8, AF4), plus to sensors (P3 and P4) which work as ground reference point and feed-forward reference to help reducing external electrical interference, respectively.

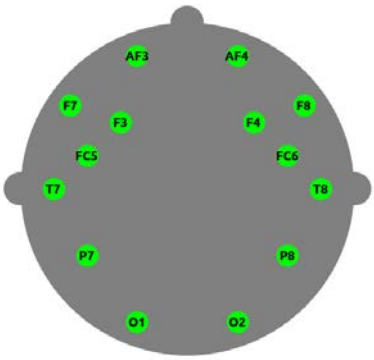


Figure 5.3: Emotiv EPOC neuroheadset sensor positions.

The Emotiv EPOC neuroheadset uses felt-based sensors with gold-plated contacts, which need to be soaked in saline solution before connecting each sensor to the scalp (to improve the electrical conductivity). Its corresponding sampling rate is 128 Hz (downsampled from 2048 Hz internally) and its effective bandwidth is 0.2-45 Hz[5].

The low-resolution recordings and artifacts induced in a mobile setup both present significant challenges. Nevertheless, the Emotiv EPOC neuroheadset has proven to be a valid alternative to laboratory ERP systems for recording reliable ERP components, such as the P300 [Debener et al., 2012] and the mismatch negativity (MMN) [Badcock et al., 2013]. It has also been used successfully in a wide variety of applications, including BCI [Campbell et al., 2010], neurofeedback [Stopczynski, Stahlhut, Petersen, et al., 2014] as well as in the EEG inverse problem [Stopczynski, Stahlhut, Larsen, et al., 2014].

5.1.2 Mobile device

The mobile device is in charge of receiving the electrical potentials sent by the wireless EEG neuroheadset, processing such data to generate the cortical activation maps and visualizing, using a 3D brain model, the brain

[5]http://emotiv.com/eeg/download_specs.php

regions that are currently active. As mobile computing platform, in this thesis we use the Samsung Galaxy Note I smartphone (Dual-core 1.4GHz ARM cortex-A9 Processor, 1 GB RAM) and the Nexus 7 tablet (Quad-core 1.3GHz Nvidia Tegra3 Processor, 1 GB RAM), both with an Android Operating System (OS).

(a) Samsung Galaxy Note I (b) Nexus 7

Figure 5.4: Mobile devices.

In order to communicate the Emotiv EPOC neuroheadset with the smartphone/tablet, we need to use the Emotiv USB dongle, which can be connected to the mobile device using a micro USB OTG cable, as shown in Figure 5.5.

Figure 5.5: From left to right: Emotiv EPOC neuroheadset, Emotiv USB dongle, micro USB OTG cable and Nexus 7 tablet.

To interact with any electronic gadget, the mobile device requires drivers, either standard kernel modules or propietary drivers created by the vendor. In the case of the Emotiv EPOC, to interact with it, the mobile device requires the HIDRAW driver, which provides a raw interface to USB and Bluetooth Human Interface Devices (HIDs). It also requires USB host mode enabled, to power the USB bus and enumerate connected USB devices (USB host mode is supported in Android 3.1 and higher). Most desktop Linux distributions have both by default, but currently most Android mobile devices support only USB host mode out-of-the-box. Therefore, we need to build and install a custom Android kernel with the HIDRAW module enabled. The basic steps to build a custom kernel are shown in Figure 5.6 (for a more detailed overview on compiling and flashing Android kernels see [Yaghmour, 2013] and references therein). In this thesis, we built a custom kernel for the Nexus 7 tablet, which is based on the Android 4.4.2 factory image[6] (this is the latest Android OS version in the time of writing this thesis). For the Samsung Galaxy Note I, we used the custom kernel image provided by the Smartphone Brain Scanner project[7].

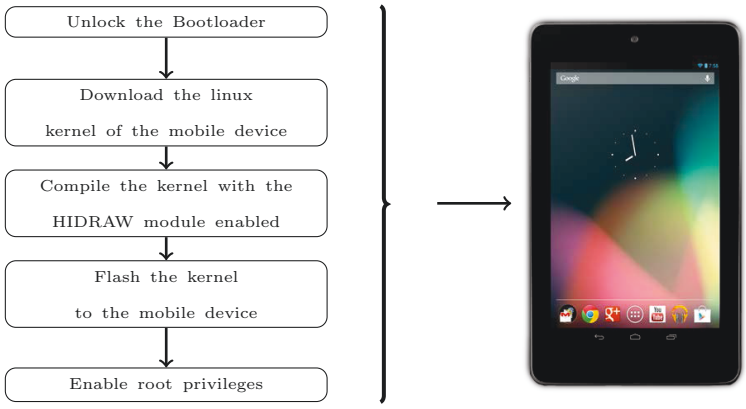


Figure 5.6: Low level prerequisites to use the Emotiv EPOC neuroheadset on Android devices.

[6]https://developers.google.com/android/nexus/images?hl=es

[7]https://github.com/SmartphoneBrainScanner/smartphonebrainscanner2-core/wiki/Downloads

5.2 Mobile brain scanner: software framework

A mobile brain scanner requires a software framework, which eases to program the interaction between the wireless EEG neuroheadset and the mobile device. Using this software framework, we could access the raw EEG data sent by the wireless EEG neuroheadset, we could apply digital signal processing methods to the raw EEG data (e.g., filtering data), we could apply an EEG inverse solver to the raw EEG data in order to generate 3D cortical activation maps, which would be visualized on the mobile device, showing us the brain regions that are currently active. As we have seen previously, the main task of a mobile brain scanner is to scan the brain and to estimate the active sources whose activity is producing the recorded EEG data. Therefore, the most important component of the software framework is the EEG inverse solver, because it determines how fast is the scanning process and how reliable are the active sources that it estimates. According to this, the component that makes the main difference between two mobile brain scanners is the EEG inverse solver that they use.

The software framework used in this thesis is the Smartphone Brain Scanner framework[8] (SBS2) [Stopczynski et al., 2011]. SBS2 is aimed for a modular framework, allowing for adding and modifying data acquisition and processing blocks. The modules are created as C++ classes and integrated directly with the core of the framework [Stopczynski, Stahlhut, Petersen, et al., 2014]. SBS2 allows to work directly with the raw EEG data, which can be recorded, including timestamped events (stimuli onsets, user response, etc). It has several signal processing methods, including filtering, and the Fast Fourier Transform, among others [Stopczynski, Stahlhut, Larsen, et al., 2014]. The EEG inverse solver currently implemented in the SBS2 framework is the LORETA EEG inverse solver [R. D. Pascual-Marqui et al., 1994]. In this thesis we propose and develop a mobile brain scanner based on the matrix factorization EEG inverse solver described in Chapter 4. The implementation details of the proposed mobile brain scanner will be explained in the following chapter.

SBS2 is written in Qt[9], a cross-platform application and UI framework for developers using C++ or QML (a CSS-JavaScript like language). Qt is currently supported on the main desktop operating systems (Linux, Mac OS, Windows), as well as on mobile devices (Android, Blackberry10, iOS and partially on Windows phones). This enables us to program one SBS2

[8] https://github.com/SmartphoneBrainScanner
[9] http://qt-project.org/

application and compile and use the same application on different target platforms.

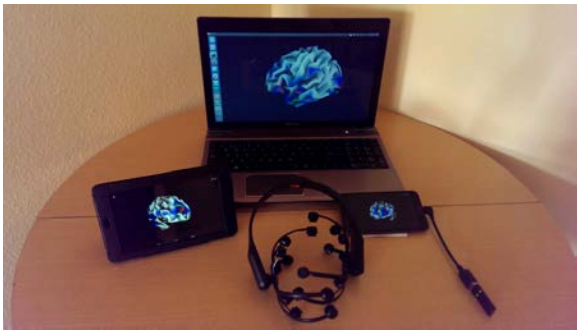


Figure 5.7: Running SBS2 on multiple target platforms.

Chapter 6

Development of a mobile brain scanner using the matrix factorization approach

As we have seen in Chapter 5, the main component of a mobile brain scanner is the EEG inverse solver, which is in charge of estimating the active sources from the recorded EEG data. In this chapter, we address the development of a mobile brain scanner based on the matrix factorization approach to solve the EEG inverse problem, which was described in Chapter 4. In this chapter we explain how to adapt the batch matrix factorization approach into an online approach, such that it can be used in a mobile brain scanner to estimate the active sources as fast as possible. We also show an experimental validation of the proposed approach, using the right index finger tapping experiment.

6.1 Implementing an online version of the matrix factorization approach

The matrix factorization approach, described in Chapter 4, works in batch mode: it receives the whole EEG matrix $\mathbf{Y}$, collected during a complete EEG experiment, and then estimates the sources matrix $\mathbf{S}$ using the alternating optimization algorithm 2. All this process is done in offline mode, this way, the time to compute the best parameters of the algorithm (the regularization

parameter (λ), the estimated rank (K) and the matrices $\mathbf{B}$ and $\mathbf{C}$) is not a concern. Moreover, we do not have any restriction on the computational resources used to run the algorithm, we can use cross-validation on a desktop machine or in a High Performance Computing (HPC) cluster to find the best parameters of the algorithm. However, as soon as we move to a mobile real-time scenario, like the one established by a mobile brain scanner, both, the time to compute the best parameters, as well as the computational resources, become into main restrictions to take into account: we receive EEG data continuously, and we need to be able to estimate and visualize the active sources as fast as possible, all this in a resources restricted device such as a smartphone or a tablet.

6.1.1 Selecting the dimensions of Y and S

In order to adapt the batch matrix factorization approach into an online approach, such that it can be used in a mobile brain scanner to estimate the active sources as fast as possible, we need to work with small matrices $\mathbf{Y}$ and $\mathbf{S}$, whose dimensions need to be selected properly. The EEG matrix $\mathbf{Y}$ has 14 rows (the number of EEG sensors of the Emotiv EPOC neuroheadset) and T columns, which correspond to the number of EEG packets to collect (in this context, we call EEG packet to a column vector (14×1) whose i-th entry corresponds to the electrical potential measured by the i-th EEG sensor). On the other hand, the sources matrix $\mathbf{S}$ has N rows (the number of sources to work with) and T columns. In the matrix factorization approach, the sources matrix is decomposed as the product $\mathbf{S} = \mathbf{BC}$, where $\mathbf{B} \in \mathbb{R}^{N \times K}$ and $\mathbf{C} \in \mathbb{R}^{K \times T}$.

To decide the number of EEG packets to collect (T) and the number of sources to work with (N), we analyzed the time required to estimate the matrices $\mathbf{B}$ and $\mathbf{C}$, each time using a different value for T and N ($T \in \{2, 4, 8\}$; $N \in \{1028, 2028\}$; and fixing $K = T$ to estimate $\mathbf{C}$, which involved finding the inverse of a $K \times K$ matrix, as fast as possible). The results can be seen in Figures 6.1, 6.2, 6.3 and 6.4. In each one of these figures we can see the mean time plus/minus one standard deviation, required to estimate the corresponding matrix using a Nexus 7 tablet, a Samsung Galaxy Note I smartphone and a Laptop (Intel Core i7-3630QM CPU at 2.4GHz, running Ubuntu 13.04). The average time was obtained using 100 repetitions.

From Figures 6.1, 6.2, 6.3 and 6.4, we can see that the fastest times were obtained, as expected, using the Laptop. With respect to the mobile devices, we can see that the Samsung Galaxy Note I was a little bit faster

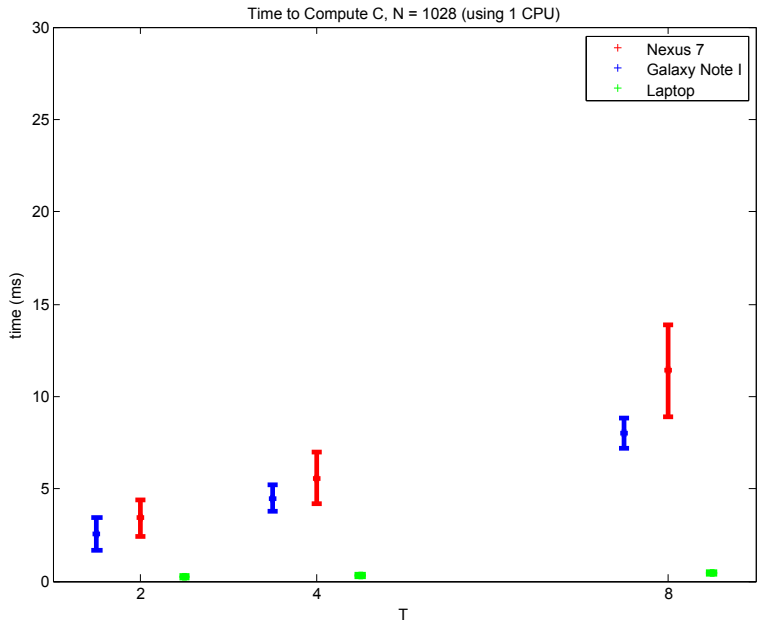


Figure 6.1: Average time to estimate $\mathbf{C}$, using $N = 1028$ sources.

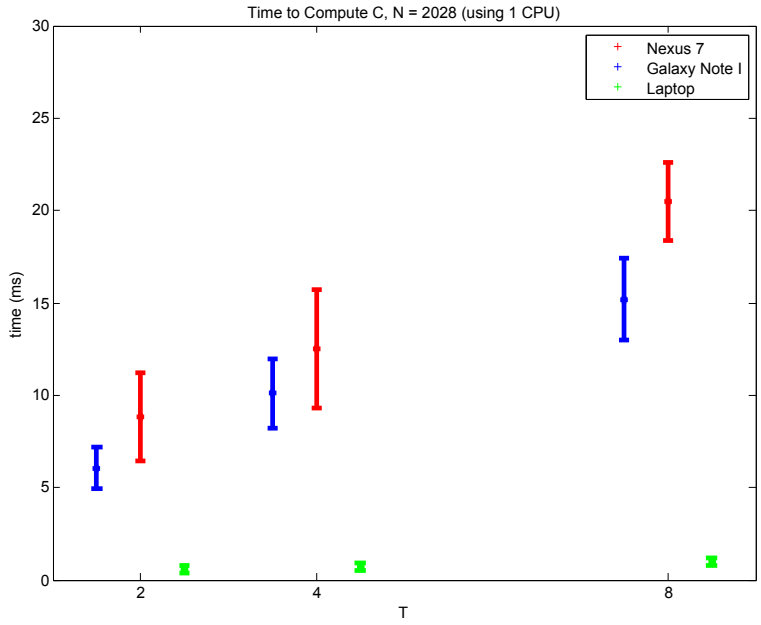


Figure 6.2: Average time to estimate $\mathbf{C}$, using $N = 2028$ sources.

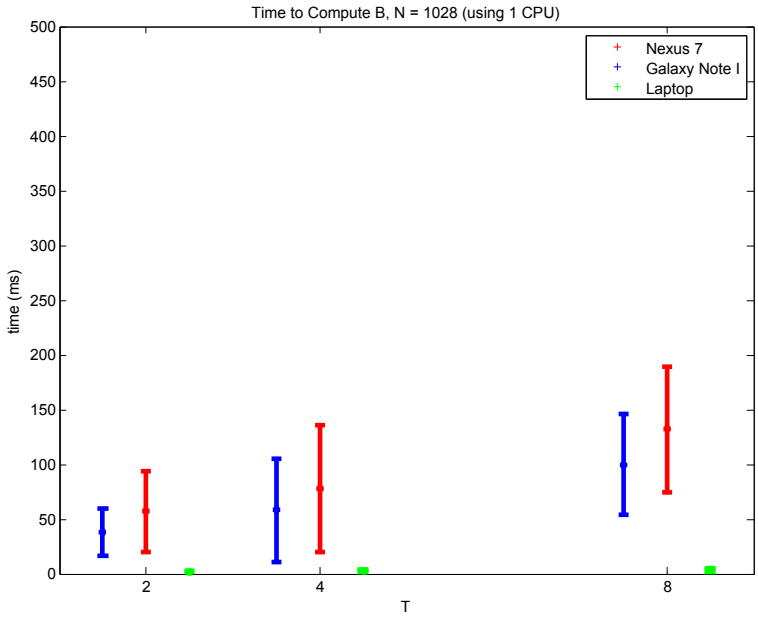


Figure 6.3: Average time to estimate $\mathbf{B}$, using $N = 1028$ sources.

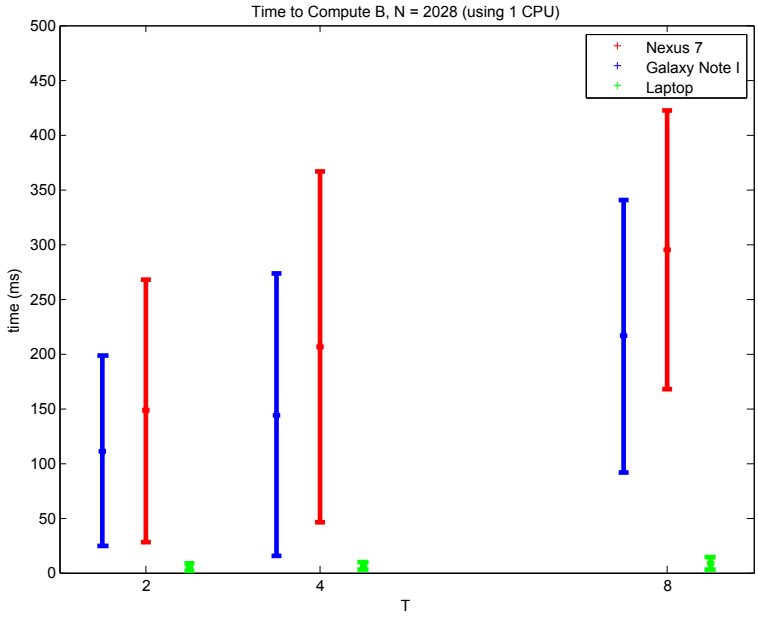


Figure 6.4: Average time to estimate $\mathbf{B}$, using $N = 2028$ sources.

than the Nexus 7. This result, which seems strange at first sight (given the computational resources of each device), can, in fact, be explained taking into account that each time we estimated $\mathbf{B}$ and $\mathbf{C}$, we used only one CPU per estimation task, and each one of the Samsung Galaxy Note I cores runs at 1.4GHz, while each one of the Nexus 7 cores runs at 1.3GHz, but they are limited to 1.2GHz when there is more than one core in use. These figures also show that the most time consuming task is the estimation of $\mathbf{B}$, while the estimation of $\mathbf{C}$ can be done really fast. By analyzing these results, we decided to work with $N = 1028$ sources, collecting $T = K = 4$ EEG packets, that is, we decided to work with matrices of the following dimensions: $\mathbf{Y} \in \mathbb{R}^{14\times 4}$ and $\mathbf{S} \in \mathbb{R}^{1028\times 4}$. The way we store the received EEG packets to create the matrix $\mathbf{Y}$ is shown in Figure 6.5. Using these values for N, T and K, according to Figures 6.1 and 6.3, we can estimate (in average) the matrix $\mathbf{C}$ in less that 10 ms and the matrix $\mathbf{B}$ in less that 100 ms (to accelerate the convergence of the FISTA algorithm used to estimate $\mathbf{B}$ (explained in Section 3.4.2), we use warm-restart, taking the previous value of $\mathbf{B}$ as a starting point for the new estimation).

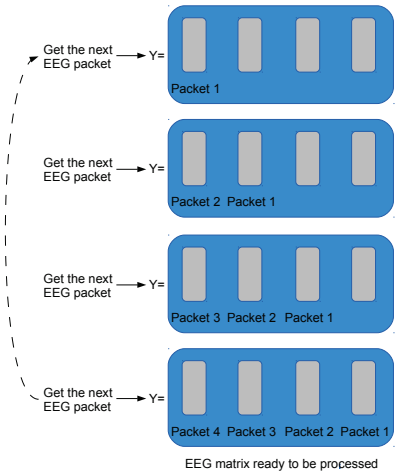


Figure 6.5: Storing EEG packets: every time a new EEG packet is received, we move the columns of $\mathbf{Y}$ to the right, this way we create a space in the first column for storing the new EEG packet received. Using this storing procedure, we guarantee to always have an EEG matrix ready to be processed.

6.1.2 Updating B and C

Given that the active sources' electrical activity changes over time (temporal dynamics enconded in **C**) quicker than their corresponding spatial locations (spatial dynamics enconded in **B**), as it is shown in Figure 6.6, the matrix **C** needs to be estimated more often than the matrix **B**. In order to accomplish this, we propose to compute **B** and **C** in parallel, using one thread for estimating **B** and one thread for estimating **C**. This gives as a result an online multi-threading matrix factorization approach, whose flow diagram is show in Figure 6.7. In this figure, we can see that the Callback object is in charge of coordinating all the processes: assemble the EEG matrix **Y**, receive and send the signals necessary for coordinating the threads in charge of estimating **B** and **C**, and request an update of the 3D brain model each time a new sources matrix is available.

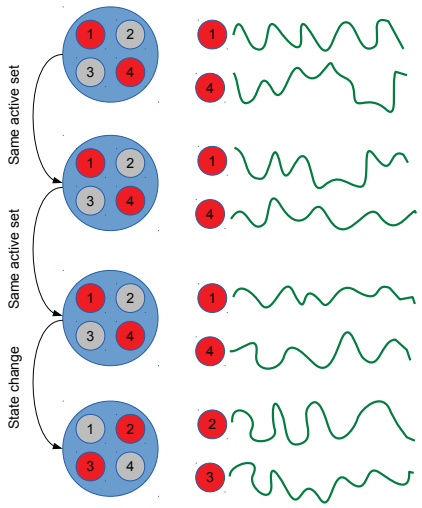


Figure 6.6: Spatial and temporal dynamics of the active sources: the set of active sources (red circles) changes slower than their corresponding electrical activity (green line).

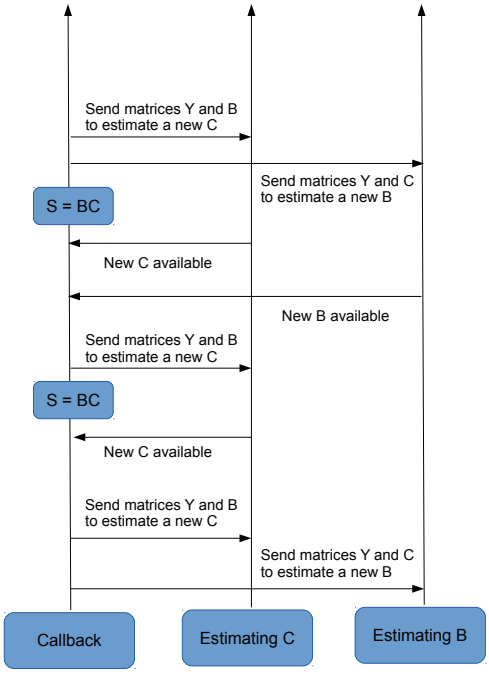


Figure 6.7: Online multi-threading matrix factorization approach.

6.1.3 Updating the regularization parameter λ

In the batch matrix factorization approach, we select λ using cross-validation, but in an online scenario this presents a big challenge:

- A big range $[\lambda_{\min},\ \lambda_{\max}]$ to analyze, would imply a lot of computations for a constrained resource device, such as a mobile device.
- It is difficult to select the appropiate range $[\lambda_{\min},\ \lambda_{\max}]$ *a priori* such that contains the optimum value of λ. In the offline scenario, we can work with a big logarithmic range in the cross-validation, and if the best λ found is equal to either $\lambda_{\min}$ or $\lambda_{\max}$, we simply increase the searching range and run the cross-validation process again in order to verify that we are obtaining the optimum value for the parameter λ, but, as we can see this is not feasible in an online scenario.

Therefore, in an online scenario would be much better to tune λ automatically as a function of the EEG data that we are processing. As

we will see in the following, this can be done by analyzing the dual function of the optimization problem associated to $\mathbf{B}$. For readibility, let us recall here the optimization problem for estimating $\mathbf{B}$:

$$\underset{\mathbf{B}}{\text{minimize}}\ \mathrm{F}(\mathbf{ABC}) + \lambda\Omega(\mathbf{B})\ , \lambda > 0 \tag{6.1}$$

where $\mathrm{F}(\mathbf{ABC}) = \frac{1}{2}\|\mathbf{ABC} - \mathbf{Y}\|_{\mathrm{F}}^2$, $\Omega(\mathbf{B}) = \|\mathbf{B}\|_{2,1}$ and $\mathbf{A}$ and $\mathbf{C}$ are fixed matrices (we can get rid of the factor $\frac{1}{2}\|\mathbf{C}\|_{\mathrm{F}}^2$ because it does not depend on $\mathbf{B}$). We can rewrite this problem as a constrained optimization problem:

$$\begin{aligned} \underset{\mathbf{B},\,\mathbf{Z}}{\text{minimize}} \quad & \mathrm{F}(\mathbf{Z}) + \lambda\Omega(\mathbf{B}) \\ \text{subject to} \quad & \mathbf{Z} = \mathbf{ABC} \end{aligned} \tag{6.2}$$

where $\mathrm{F}(\mathbf{Z}) = \frac{1}{2}\|\mathbf{Z} - \mathbf{Y}\|_{\mathrm{F}}^2$. The Lagrangian function associated to this problem is

$$L(\mathbf{B}, \mathbf{Z}, \boldsymbol{\nu}) = \mathrm{F}(\mathbf{Z}) + \Omega(\mathbf{B}) + \langle \boldsymbol{\nu}, \mathbf{ABC} - \mathbf{Z} \rangle$$

where $\langle \mathbf{P}, \mathbf{Q} \rangle$ denotes the standard inner product for matrices:

$$\langle \mathbf{P}, \mathbf{Q} \rangle = \mathrm{trace}(\mathbf{P}^\top \mathbf{Q}),\ \mathbf{P}, \mathbf{Q} \in \mathbb{R}^{M \times N}$$

To obtain the dual function $g(\boldsymbol{\nu})$ associated to (6.2), we need to find the infimum of $L(\mathbf{B}, \mathbf{Z}, \boldsymbol{\nu})$ with respect to $\mathbf{B}$ and $\mathbf{Z}$:

$$\begin{aligned} g(\boldsymbol{\nu}) &= \inf_{\mathbf{B},\,\mathbf{Z}} L(\mathbf{B}, \mathbf{Z}, \boldsymbol{\nu}) \\ &= \inf_{\mathbf{Z}} \left\{ -\left(\langle \boldsymbol{\nu}, \mathbf{Z} \rangle - \mathrm{F}(\mathbf{Z}) \right) \right\} + \inf_{\mathbf{B}} \left\{ -\left(\langle -\mathbf{A}^\top \boldsymbol{\nu} \mathbf{C}^\top, \mathbf{B} \rangle - \lambda\Omega(\mathbf{B}) \right) \right\} \\ &= -\sup_{\mathbf{Z}} \left\{ \left(\langle \boldsymbol{\nu}, \mathbf{Z} \rangle - \mathrm{F}(\mathbf{Z}) \right) \right\} - \sup_{\mathbf{B}} \left\{ \left(\langle -\mathbf{A}^\top \boldsymbol{\nu} \mathbf{C}^\top, \mathbf{B} \rangle - \lambda\Omega(\mathbf{B}) \right) \right\} \\ &= -\mathrm{F}^*(\boldsymbol{\nu}) - \lambda\Omega^*\left(-\frac{1}{\lambda} \mathbf{A}^\top \boldsymbol{\nu} \mathbf{C}^\top \right) \end{aligned} \tag{6.3}$$

where F^* and Ω^* denote the conjugate function of F and Ω, respectively:

$$\begin{aligned} \mathrm{F}^*(\boldsymbol{\nu}) &= \frac{1}{2}\|\boldsymbol{\nu}\|_{\mathrm{F}}^2 + \mathrm{trace}(\boldsymbol{\nu}^\top \mathbf{Y}) \\ \Omega^*\left(-\frac{1}{\lambda} \mathbf{A}^\top \boldsymbol{\nu} \mathbf{C}^\top \right) &= \begin{cases} 0, & \text{if } \|-\frac{1}{\lambda}\mathbf{A}^\top \boldsymbol{\nu} \mathbf{C}^\top\|_{2,\infty} \leq 1 \\ +\infty, & \text{otherwise} \end{cases} \end{aligned}$$

Note that $\Omega^*\left(-\frac{1}{\lambda}\mathbf{A}^\top \boldsymbol{\nu} \mathbf{C}^\top\right)$ is the indicator function of the unit-ball associated to the dual norm of Ω (the dual norm associated to $\|\cdot\|_{2,1}$ is $\|\cdot\|_{2,\infty}$).

The primal problem (6.2) is convex and satisfies the Slater's conditions, therefore strong duality holds and the KKT conditions also holds:

$$\begin{aligned} \mathbf{ABC} &= \mathbf{Z} && (6.4)\\ \nabla L_{\mathbf{Z}}(\mathbf{Z}, \boldsymbol{\nu}) &= \mathbf{0} && (6.5)\\ \mathbf{0} &\in \partial\{L_{\mathbf{B}}(\mathbf{B}, \boldsymbol{\nu})\} && (6.6) \end{aligned}$$

where $L_{\mathbf{Z}}(\mathbf{Z}, \boldsymbol{\nu}) = \mathrm{F}(\mathbf{Z}) - \langle \boldsymbol{\nu}, \mathbf{Z} \rangle$ and $L_{\mathbf{B}}(\mathbf{B}, \boldsymbol{\nu}) = \langle \mathbf{A}^\top \boldsymbol{\nu} \mathbf{C}^\top, \mathbf{B} \rangle + \lambda \Omega(\mathbf{B})$. Using (6.5) we get the following:

$$\begin{aligned} \nabla L_{\mathbf{Z}}(\mathbf{Z}, \boldsymbol{\nu}) = \nabla \mathrm{F}(\mathbf{Z}) - \boldsymbol{\nu} &= \mathbf{0} \\ &\Rightarrow \boldsymbol{\nu} = \nabla \mathrm{F}(\mathbf{Z}) = \mathbf{Z} - \mathbf{Y} \\ &\quad \boldsymbol{\nu} = \mathbf{ABC} - \mathbf{Y} && (6.7) \end{aligned}$$

We need to guarantee that $g(\boldsymbol{\nu})$ is finite. In order to guarantee this, we use the condition that appears in the definition of $\Omega^*\left(-\frac{1}{\lambda}\mathbf{A}^\top \boldsymbol{\nu} \mathbf{C}^\top\right)$, in combination with (6.7):

$$\begin{aligned} \|-\frac{1}{\lambda}\mathbf{A}^\top \boldsymbol{\nu} \mathbf{C}^\top\|_{2,\infty} &\leq 1 \\ \|\mathbf{A}^\top (\mathbf{ABC} - \mathbf{Y}) \mathbf{C}^\top\|_{2,\infty} &\leq \lambda \end{aligned}$$

Let us denote by $\lambda_{\max}$ the value of λ that yields the null solution $\mathbf{B} = \mathbf{0}$, hence:

$$\|\mathbf{A}^\top \mathbf{Y} \mathbf{C}^\top\|_{2,\infty} = \lambda_{\max} \tag{6.8}$$

Note that using any value of λ greater than $\lambda_{\max}$ will yield the null solution $\mathbf{B} = \mathbf{0}$. As we can see, $\lambda_{\max}$ depends on the EEG matrix $\mathbf{Y}$, as well as on the matrices $\mathbf{A}$ and $\mathbf{C}$, therefore, it could be used to tune λ. We propose to tune λ as follows:

$$\lambda = \text{factor} \cdot \lambda_{\max} \tag{6.9}$$

where factor $\in (0, 1)$ is a scalar that the user should specify, and denotes the percentage of the $\lambda_{\max}$ to be used as regularization parameter. Note that (6.9) should be evaluated before running FISTA for estimating $\mathbf{B}$. This way, every time we need to update $\mathbf{B}$, we will use a different regularization parameter that depends on the geometry of the problem (the lead-field matrix $\mathbf{A}$), the current EEG matrix that we are processing ($\mathbf{Y}$), as well as the current estimation of the temporal dynamics of the active sources ($\mathbf{C}$).

Using the online multi-threading matrix factorization approach showed in Figure 6.7, together with (6.9) for updating the regularization parameter

λ, we can update the temporal and spatial dynamics of the active sources several times a second. Figures 6.8 and 6.9 show a snapshot of the Samsung Galaxy Note I and the Nexus 7 while running the proposed online multi-threading matrix factorization EEG inverse solver.

Figure 6.8: Online matrix factorization solver running on the Samsung Galaxy Note I (gray = no activity, blue = low activity, green = medium activity, red = high activity).

Figure 6.9: Online matrix factorization solver running on the Nexus 7.

In order to implement this solver as a C++ module, we use the Eigen library [Guennebaud et al., 2010], which is a C++ template library for linear algebra. This library supports all the matrix sizes, all the standard numeric types and includes various matrix decompositions. It performs explicit

vectorization for SSE 2/3/4, ARM NEON and AltiVec instruction sets, and fixed-sized matrices are fully optimized: dynamic memory allocation is avoided and the loops are unrolled when that makes sense. This library is very friendly to use with the Android OS, it does not have any dependencies other than the C++ standard library. Besides, Eigen is a pure template library defined in the headers, therefore there is no binary library to link to.

6.2 Experimental validation: right index finger tapping

The exact validation of EEG inverse methods is not possible due to the lack of a ground truth. Therefore, a standard way of evaluating inverse methods is to assess their ability to estimate known activity. This is done here by estimating the electrical activity of the active sources related to the right index finger tapping experiment. It is known that the voluntary movement of one finger of the right hand leads to a suppression in the electrical activity of the alpha band (8-13 Hz) over the left pre-motor region. Hence, in this section, we use the online matrix factorization approach to estimate the electrical activity of the active sources that belong to the left pre-motor region, in order to verify if our approach finds the expected suppression in the electrical activity. We compare the performance of the online matrix factorization approach against the performance of the LORETA EEG inverse solver (explained in Section 3.3.2), which is the one used in the standard mobile brain scanner currently implemented in the SBS2 framework. We also compare, visually, the position of the active sources found in the left pre-motor region using fMRI [Taniguchi et al., 2000] against the ones found using the online matrix factorization and the LORETA EEG inverse solvers, respectively.

6.2.1 Experimental setup

The subject sat comfortably in front of a table, with his right hand resting on the table. The hand was positioned palm down, so that the index finger could be moved to tap on the table, as it is shown in Figure 6.10. A motor sequence is defined by a successive tapping of the index finger. Visual stimuli were presented at a constant viewing distance, and consisted of a "X" to indicate the starting of the motor sequence. The subject was asked to continue tapping until the word "Relax" appears on the screen. We presented 100 times the visual stimuli.

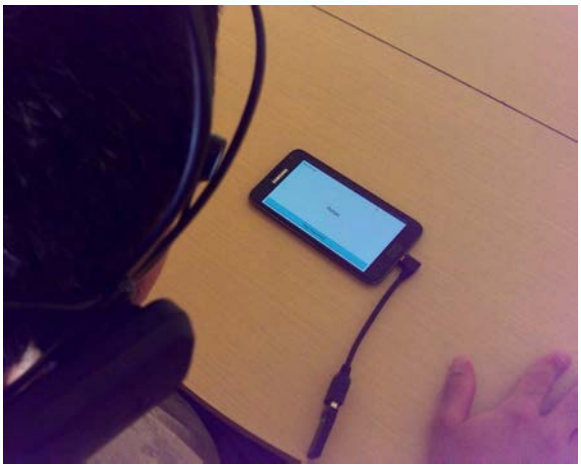


Figure 6.10: Finger tapping experiment using the Samsung Galaxy Note I as recording device.

6.2.2 Performance evaluation

Figure 6.11 shows how the alpha activity is suppressed over time in the left pre-motor region, which corresponds to the Pre-central Left AAL region of the Automated Anatomical Labeling (AAL) brain atlas [Tzourio-Mazoyer et al., 2002]. This demostrates the ability of the online matrix factorization approach and the LORETA EEG inverse solver to reconstruct meaningful activity within the given region. However, the suppression, as well as the recovery, is more pronounced in the alpha activity estimated using the online matrix factorization approach. This rapid and pronounced change of state, suppression-recovery, is the expected behavior, according to [Pfurtscheller & Lopes da Silva, 1999]. This figure also shows that the alpha activity is suppressed as soon as the finger starts to move. Both responses shown in the Figure 6.11 were generated as the average response over the 100 trials of the motor sequence. Note that, while the result is presented as an average over runs, the sources localization was carried out in online mode with model parameters and sources matrix estimated online.

Figures 6.12 and 6.13 show the localization of the active sources in the Pre-central Left AAL region during the right index finger tapping. As we can see from Figure 6.13, LORETA does not promote sparse solutions, and according to it all the sources in the Pre-central Left AAL region are active during the right index finger tapping. Unlike LORETA, the online matrix factorization approach obtains a sparse and focal solution. By comparing

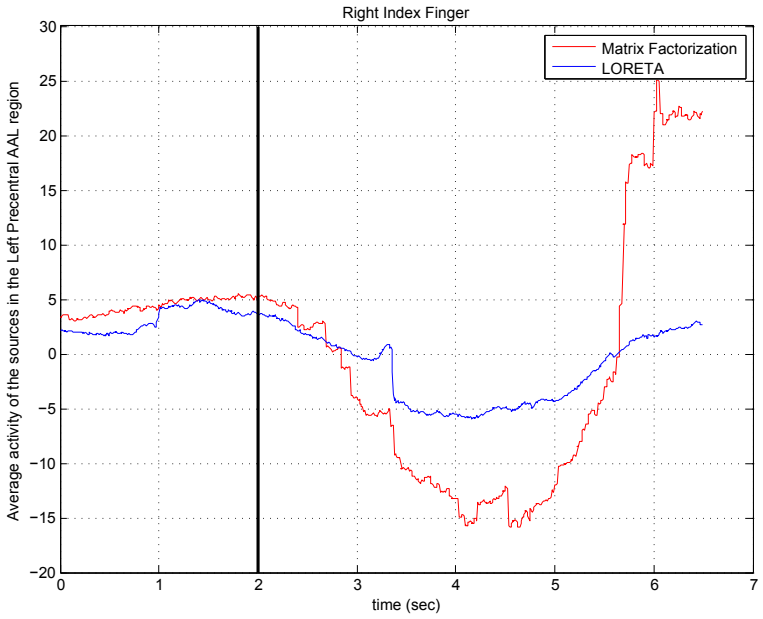
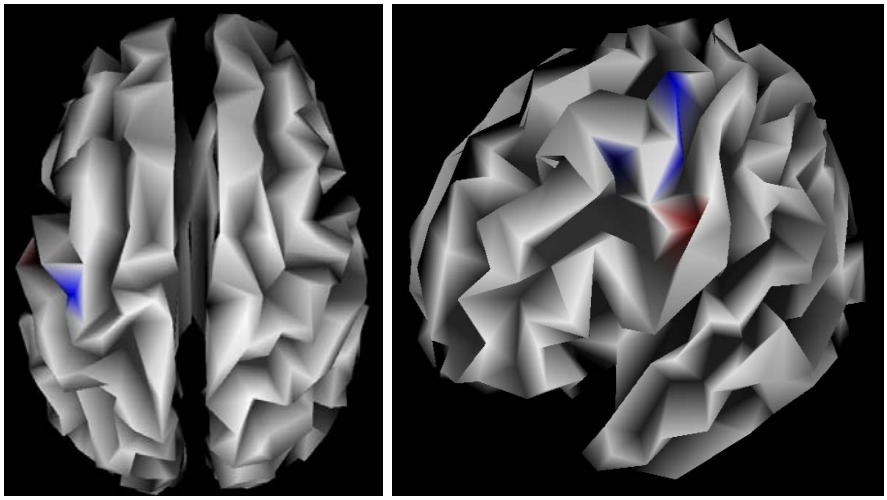


Figure 6.11: Right index finger tapping experiment using the Emotiv EPOC neuroheadset. The black line indicates the starting of the finger movement.

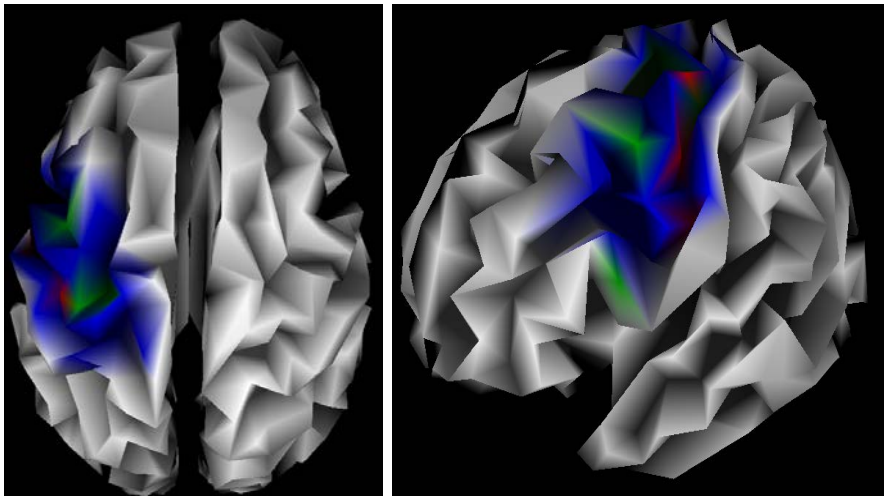
Figures 6.12 and 6.13 with Figure 6.14, which shows the fMRI statistical image of power changes in the alpha band during the right index finger movement, we can see that both methods found active sources in the area highlighted by the fMRI image. Nevertheless, the solution found by the online matrix factorization approach is as focal as the active region shown in the statistical image, unlike the solution found by LORETA, which is too diffuse, hence proving the efectiveness of the proposed approach.



(a) Axial view of the Pre-central Left AAL region

(b) Sagittal view of the Pre-central Left AAL region

Figure 6.12: Active sources estimated by the online matrix factorization approach.



(a) Axial view of the Pre-central Left AAL region

(b) Sagittal view of the Pre-central Left AAL region

Figure 6.13: Active sources estimated by the LORETA EEG inverse solver.

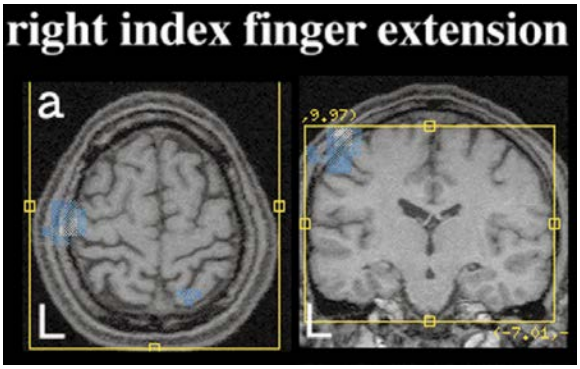


Figure 6.14: fMRI statistical image of power changes in the alpha band during the right index finger movement (adapted from [Taniguchi et al., 2000]).

Chapter 7

Conclusion and future work

7.1 Conclusion

In this thesis we have addressed the development of a mobile brain scanner, which is based on a wireless EEG neuroheadset, in charge of acquiring and transmitting the electrical potential measured on the scalp, and one mobile device (smartphone or tablet), in charge of receiving and processing these data in order to estimate and show, using a 3D brain model, the cortical areas that are currently active.

Unlike fMRI, this mobile brain scanner has direct access to the brain electrical activity, hence it is not susceptible to influence by non-neural changes in the body; it has high temporal resolution, which enable to scan the brain on the millisecond timescale on which neurons operate; and it can be used on a wide variety of scenarios, in which the person being scanned can move and act normally, without being restricted to a tiny closed space (like the MRI scanner), which enable to study the brain under naturalistic conditions.

To estimate the cortical areas that are currently active, the mobile brain scanner needs to solve an electromagnetic inverse problem called the EEG inverse problem. This problem is underdetermined, ill-posed and has infinite solutions, mainly because of the small number of EEG electrodes available to capture the electrical activity produced by thousands of brain current sources and the low spatial resolution of the EEG caused by the low conductivity of the skull, which blurs the electrical potential on the scalp. To solve this problem, in this thesis we have proposed a new method based on regularization theory, which involves the replacement of the original problem with a nearby well-posed problem whose solution approximates the

required solution. To reduce the complexity of this problem, in this thesis we assumed that the number of active sources is small, that is, we assumed that the set of active sources is a sparse set. Additionally, we also assumed a linear relationship between the electrical activity of the main active sources and the electrical activity that they induce on the surrounding neighbors. The proposed method is able to handle, explicitly, the spatial and temporal dynamics of the active sources by decomposing the sources matrix into the product of two matrices, each of which encodes the spatial and temporal behavior of the active sources, respectively.

To develop the mobile brain scanner, in this thesis we have used the Smartphone Brain Scanner software framework (SBS2), which is a software platform that provides the basic infrastructure for developing a mobile brain scanner. SBS2 eases to program the interaction between the wireless EEG neuroheadset and the mobile device. As a wireless EEG neuroheadset, we have used the Emotiv EPOC neuroheadset, which has 14 EEG electrodes and a sampling rate of 128 Hz. Finally, as a mobile device, we have used the Samsung Galaxy Note I smartphone and the Nexus 7 tablet, both running an Android Operating System.

To summarize:

- We have developed a novel approach to solve the EEG inverse problem, which takes into account the structured sparsity and low rank of the sources matrix.

  - The method is based on the factorization of the sources matrix as a product of two matrices: the first one encodes the spatial dynamics of the sources (how they change their activation patterns), while the second one encodes their corresponding temporal dynamics (how they change their electrical activity over time). Our method combines the ideas of the Group Lasso (structured sparsity) and Trace Norm (low rank) into one unified framework.
  - We have also developed and analyzed the convergence of an alternating minimization algorithm to solve the resulting nonsmooth-nonconvex regularization problem.
  - We brought our method to a real life scenario: online solving of the EEG inverse problem on a mobile device, which is continuously supplied with EEG data coming from a wireless EEG neuroheadset.

- We have developed a mobile brain scanner, which is able to show, online and using a 3D brain model, the cortical areas that are currently active.

  - To solve the EEG inverse problem in the mobile brain scanner, we developed an online multi-threading version of the matrix factorization approach, and implemented it as a C++ multiplatform module that can be used on mobile devices as well as on desktop platforms.
  - We contributed to extend the SBS2 framework with the following additions:
    * An EEG inverse solver based on the online multi-threading version of the matrix factorization approach.
    * An EEG inverse solver based on the Lasso regularizer.
    * An EEG inverse solver based on the Group Lasso regularizer.
    * We have updated the SBS2 framework to Qt5.3, which is the lastest version of the Qt framework in the time of writing this thesis.
    * We have added a new 3D brain visualization class, which is based on the new QWindow and QOpenGL* classes.

    The code is available at:

    https://github.com/jmontoyam/SBS2_Qt5_and_SSO
  - We contributed to provide to the EEG, machine learning and signal processing communities with a new C++ template library for Structured Sparse Optimization, which can be used on mobile devices (currently tested on Android OS), as well as on desktop platforms. The code is available at: https://github.com/jmontoyam/SSO

- The research work done during this thesis has generated the following publications:

  - **Montoya-Martínez, J.**, Artés-Rodríguez, A., Pontil, M. and Hansen, L.K. "A regularized matrix factorization approach to induce structured sparse-low-rank solutions in the EEG inverse problem." *EURASIP Journal on Advances in Signal Processing* 2014, **2014**:97.
  - **Montoya-Martínez, J.**, Artés-Rodríguez, A. and Pontil, M. "Structured Sparse-Low Rank Matrix Factorization for the EEG

Inverse Problem." *Proceedings of the 4th International Workshop on Cognitive Information Processing (CIP)*, Copenhagen, Denmark, 2014.

– **Montoya-Martínez, J.**, Artés-Rodríguez, A., Hansen, L.K and Pontil, M. "Structured sparsity regularization approach to the EEG inverse problems." *Proceedings of the 3rd International Workshop on Cognitive Information Processing (CIP)*, Baiona, Spain, 2012.

7.2 Future work

As a continuation of the research work that we have done in this thesis, we propose to incorporate new neurophysiological prior knowledge into the formulation of the EEG inverse problem, in order to improve the estimation of the active sources, as well as their corresponding electrical activity. We also want to use this new prior knowledge to solve the EEG inverse problem in such a way that the resulting solution helps the researcher to answer the following questions:

1. Which are the most active brain areas related with the processing of a given stimulus?

   To answer this question, we propose to group the sources that belong to the same brain area into one single group. Then, we can use a regularization function that induces sparsity inside the set of all Brain Areas (BA), such as the following one:

   $$\Omega(\{\text{set of all BA}\}) = \sum_{i=1}^{|BA|} \|BA_i\|_2 \tag{7.1}$$

   where $|BA|$ and BA_i denote the cardinality of the set of all brain areas and the i-th brain area, respectively.

2. Which are the most active brain areas, as well as the most active sources inside them, related with the processing of a given stimulus?

   In order to answer this question, we propose to use a regularization function that induces sparsity in the set of all BA, as well as inside each one of them. This can be done by using the following regularizer:

   $$\Omega(\{\text{set of all BA}\}) = \sum_{i=1}^{|BA|} \left(\|BA_i\|_2 + \sum_{j \in BA_i} \|\mathbf{B}_i^j\|_2 \right) \tag{7.2}$$

where $\mathbf{B}_i^j$ denotes the row of the coding matrix $\mathbf{B}$ that corresponds to the j-th source, which belongs to the i-th brain area (the matrix $\mathbf{B}$ encodes the spatial dynamics of the sources).

These brain areas could be taken from a brain atlas that creates the areas based on anatomical features or functional connectivity of the brain. Usually, these brain areas could be overlapped. When this happens, the proximity operators associated to the regularizers (7.1) and (7.2) may not be easily computable, which could degrade the performance of the optimization algorithm used to estimate the position of the active sources (FISTA). To tackle this scenario, we propose to use other optimization methods that can efficiently handle this overlapping, such as the Alternating Direction Method of Multipliers (ADMM) [Hestenes, 1969; Powell, 1969; D. Bertsekas, 1996; S. Boyd et al., 2011] or the fixed point based method proposed in [Argyriou et al., 2011], which computes the proximity operator of a composite regularizer from the solution of a certain fixed point problem, which depends on the structure of the function Ω as well as the corresponding overlapping pattern.

References

Abernethy, J., Bach, F., Evgeniou, T., & Vert, J.-P. (2009). A new approach to collaborative filtering: Operator estimation with spectral regularization. *The Journal of Machine Learning Research*, *10*, 803–826.

Andrew, C., & Pfurtscheller, G. (1996). Event-related coherence as a tool for studying dynamic interaction of brain regions. *Electroencephalography and clinical Neurophysiology*, *98*(2), 144–148.

Argyriou, A., Evgeniou, T., & Pontil, M. (2007). Multi-task feature learning. *Advances in neural information processing systems*, *19*, 41–48.

Argyriou, A., Evgeniou, T., & Pontil, M. (2008). Convex multi-task feature learning. *Machine Learning*, *73*(3), 243–272.

Argyriou, A., Micchelli, C., Pontil, M., Shen, L., & Xu, Y. (2011). Efficient first order methods for linear composite regularizers. *Arxiv preprint arXiv:1104.1436*.

Bach, F., Jenatton, R., Mairal, J., & Obozinski, G. (2011). Optimization with Sparsity-Inducing Penalties. *Foundations and Trends in Machine Learning*, *4*(1), 1–106.

Badcock, N. A., Mousikou, P., Mahajan, Y., de Lissa, P., Thie, J., & McArthur, G. (2013). Validation of the Emotiv EPOC® EEG gaming system for measuring research quality auditory ERPs. *PeerJ 1:e38 http://dx.doi.org/10.7717/peerj.38*.

Bear, M. F., Connors, B. W., & Paradiso, M. A. (2007). *Neuroscience* (Vol. 2). Lippincott Williams & Wilkins.

Beck, A., & Teboulle, M. (2009). A fast iterative shrinkage-thresholding algorithm for linear inverse problems. *SIAM Journal on Imaging Sciences*, *2*(1), 183–202.

Bentin, S., Allison, T., Puce, A., Perez, E., & McCarthy, G. (1996). Electrophysiological studies of face perception in humans. *Journal of cognitive neuroscience*, *8*(6), 551–565.

Bertsekas, D. (1996). *Constrained optimization and Lagrange multiplier methods*. Athena Scientific.

Bertsekas, D. (1999). *Nonlinear programming*. Athena Scientific.

Bertsekas, D. P. (2009). *Convex optimization theory*. Athena Scientific.

Blackwood, D., & Muir, W. (1990). Cognitive brain potentials and their application. *The British Journal of Psychiatry*, *157*(9), 96–101.

Borwein, J., Lewis, A., Borwein, J., & Lewis, A. (2006). *Convex analysis and nonlinear optimization: theory and examples*. Springer New York.

Boyd, S., Parikh, N., Chu, E., Peleato, B., & Eckstein, J. (2011). *Distributed optimization and statistical learning via the alternating direction method of multipliers*. Now Publishers.

Boyd, S. P., & Vandenberghe, L. (2004). *Convex optimization*. Cambridge University Press.

Brodmann, K. (1909). *Vergleichende Lokalisationslehre der Grosshirnrinde in ihren Prinzipien dargestellt auf Grund des Zellenbaues*. Barth.

Campbell, A., Choudhury, T., Hu, S., Lu, H., Mukerjee, M. K., Rabbi, M., & Raizada, R. D. (2010). NeuroPhone: brain-mobile phone interface using a wireless EEG headset. In *Proceedings of the second ACM SIGCOMM workshop on Networking, systems, and applications on mobile handhelds* (pp. 3–8).

Chen, S., Donoho, D., & Saunders, M. (1999). Atomic decomposition by basis pursuit. *SIAM journal on scientific computing*, *20*(1), 33–61.

Collura, T. F. (1993). History and evolution of electroencephalographic instruments and techniques. *Journal of clinical neurophysiology*, *10*(4), 476–504.

Combettes, P. L., & Pesquet, J.-C. (2011). Proximal splitting methods in signal processing. In *Fixed-point algorithms for inverse problems in science and engineering* (pp. 185–212). Springer.

Combettes, P. L., & Wajs, V. R. (2005). Signal recovery by proximal forward-backward splitting. *Multiscale Modeling & Simulation*, *4*(4), 1168–1200.

Debener, S., Minow, F., Emkes, R., Gandras, K., & Vos, M. (2012). How about taking a low-cost, small, and wireless EEG for a walk? *Psychophysiology*, *49*(11), 1617–1621.

Deecke, L., & Kornhuber, H. H. (1978). An electrical sign of participation of the mesial 'supplementary' motor cortex in human voluntary finger movement. *Brain research*, *159*(2), 473–476.

Donoho, D. L., & Johnstone, I. M. (1995). Adapting to unknown smoothness via wavelet shrinkage. *Journal of the american statistical association*, *90*(432), 1200–1224.

Frank, E. (2004). Electric potential produced by two point current sources in a homogeneous conducting sphere. *Journal of applied physics*, *23*(11), 1225–1228.

Friedman, J., Hastie, T., & Tibshirani, R. (2001). *The elements of statistical learning* (Vol. 1). Springer Series in Statistics New York.

Friedman, J., Hastie, T., & Tibshirani, R. (2010). A note on the group lasso and a sparse group lasso. *arXiv preprint arXiv:1001.0736*.

Friston, K., Harrison, L., Daunizeau, J., Kiebel, S., Phillips, C., Trujillo-Barreto, N., ... Mattout, J. (2008). Multiple sparse priors for the M/EEG inverse problem. *NeuroImage*, *39*(3), 1104–1120.

Gramfort, A. (2009). *Mapping, timing and tracking cortical activations with MEG and EEG: Methods and application to human vision.* Unpublished doctoral dissertation. Retrieved from `http://tel.archives-ouvertes.fr/tel-00426852/fr/`

Gramfort, A., Strohmeier, D., Haueisen, J., Hamalainen, M., & Kowalski, M. (2011). Functional brain imaging with M/EEG using structured sparsity in time-frequency dictionaries. In *Information Processing in Medical Imaging* (pp. 600–611).

Gramfort, A., Strohmeier, D., Haueisen, J., Hämäläinen, M. S., & Kowalski, M. (2013). Time-frequency mixed-norm estimates: sparse M/EEG imaging with non-stationary source activations. *NeuroImage*, *70*, 410–422.

Guennebaud, G., Jacob, B., et al. (2010). *Eigen v3.* http://eigen.tuxfamily.org.

Gunawardana, A., & Byrne, W. (2005). Convergence theorems for generalized alternating minimization procedures. *The Journal of Machine Learning Research*, *6*, 2049–2073.

Hallez, H., Vanrumste, B., Grech, R., Muscat, J., De Clercq, W., Vergult, A., ... Lemahieu, I. (2007). Review on solving the forward problem in EEG source analysis. *Journal of neuroengineering and rehabilitation*, *4*(1), 46.

Hämäläinen, M., Hari, R., Ilmoniemi, R., Knuutila, J., & Lounasmaa, O. (1993). Magnetoencephalography—theory, instrumentation, and applications to noninvasive studies of the working human brain. *Reviews of modern Physics*, *65*(2), 413.

Hämäläinen, M. S., & Ilmoniemi, R. (1994). Interpreting magnetic fields of the brain: minimum norm estimates. *Medical & biological engineering & computing*, *32*(1), 35–42.

Hansen, P. C. (1998). *Rank-deficient and discrete ill-posed problems: numerical aspects of linear inversion.* SIAM.

Haufe, S., Tomioka, R., Dickhaus, T., Sannelli, C., Blankertz, B., Nolte, G., & Müller, K. (2011). Large-scale EEG/MEG source localization with spatial flexibility. *NeuroImage*, *54*(2), 851–859.

Henson, R., Goshen-Gottstein, Y., Ganel, T., Otten, L., Quayle, A., & Rugg, M. (2003). Electrophysiological and haemodynamic correlates of face perception, recognition and priming. *Cerebral Cortex*, *13*(7), 793.

Hestenes, M. R. (1969). Multiplier and gradient methods. *Journal of optimization theory and applications*, *4*(5), 303–320.

Horn, R., & Johnson, C. (1990). *Matrix analysis.* Cambridge University Press.

Im, C.-H., & He, B. (2006). On the Accurate Anatomical Constraints in EEG/MEG Source Imaging. *International Journal of Bioelectromagnetism*, *8*(1), III/1–III/11.

Jasper, H. H. (1958). The ten twenty electrode system of the international federation. *Electroencephalography and clinical neurophysiology*, *10*, 371–375.

Jenatton, R., Mairal, J., Obozinski, G., & Bach, F. (2010). Proximal methods for sparse hierarchical dictionary learning. In *Proceedings of the International Conference on Machine Learning (ICML)*.

Kandel, E. R., Schwartz, J. H., Jessell, T. M., et al. (2000). *Principles of neural science* (Vol. 4). McGraw-Hill New York.

Kutas, M., & Hillyard, S. A. (1980). Reading between the lines: Event-related brain potentials during natural sentence processing. *Brain and language*, *11*(2), 354–373.

Litvak, V., Mattout, J., Kiebel, S., Phillips, C., Henson, R., Kilner, J., ... Friston, K. (2011). EEG and MEG data analysis in SPM8. *Computational intelligence and neuroscience, vol. 2011, Article ID 852961, 32 pages, 2011. doi:10.1155/2011/852961*.

Luck, S. J. (2005). *An introduction to the event-related potential technique*. MIT Press.

Mairal, J., Jenatton, R., Obozinski, G., & Bach, F. (2010). Network flow algorithms for structured sparsity. *Arxiv preprint arXiv:1008.5209*.

Malmivuo, J., & Plonsey, R. (1995). *Bioelectromagnetism: principles and applications of bioelectric and biomagnetic fields*. Oxford University Press.

Menendez, R. G. d. P., Murray, M., Michel, C., Martuzzi, R., & Andino, S. G. (2004). Electrical neuroimaging based on biophysical constraints. *NeuroImage*, *21*(2), 527–539.

Micchelli, C. A., Morales, J. M., & Pontil, M. (2013). Regularizers for structured sparsity. *Advances in Computational Mathematics*, *38*(3), 455–489.

Micchelli, C. A., Shen, L., & Xu, Y. (2011). Proximity algorithms for image models: denoising. *Inverse Problems*, *27*(4), 045009.

Moreau, J. (1965). Proximité et dualité dans un espace hilbertien. *Bull. Soc. Math. France*, *93*(2), 273–299.

Mosher, J. C., Leahy, R. M., & Lewis, P. S. (1999). EEG and MEG: forward solutions for inverse methods. *Biomedical Engineering, IEEE Transactions on*, *46*(3), 245–259.

Murakami, S., & Okada, Y. (2006). Contributions of principal neocortical neurons to magnetoencephalography and electroencephalography signals. *The Journal of physiology*, *575*(3), 925–936.

Natarajan, B. K. (1995). Sparse approximate solutions to linear systems. *SIAM journal on computing*, *24*(2), 227–234.

Nesterov, Y. (2007). Gradient methods for minimizing composite objective function. CORE Discussion Papers 2007076, Université catholique de Louvain. *Center for Operations Research and Econometrics (CORE)*.

Nunez, P. L., & Srinivasan, R. (2006). *Electric fields of the brain: the neurophysics of EEG*. Oxford University Press.

Ou, W., Hämäläinen, M., & Golland, P. (2009). A distributed spatio-temporal EEG/MEG inverse solver. *NeuroImage*, *44*(3), 932–946.

Pascual-Marqui, R. (1999). Review of methods for solving the EEG inverse problem. *International Journal of Bioelectromagnetism*, *1*(1), 75–86.

Pascual-Marqui, R. D., Michel, C. M., & Lehmann, D. (1994). Low resolution electromagnetic tomography: a new method for localizing electrical activity in the brain. *International Journal of psychophysiology*, *18*(1), 49–65.

Patterson, M. M. (1973). *Bioelectric recording techniques* (Vol. 1). Academic Press.

Peterson, N. N., Schroeder, C. E., & Arezzo, J. C. (1995). Neural generators of early cortical somatosensory evoked potentials in the awake monkey. *Electroencephalography and Clinical Neurophysiology/Evoked Potentials Section*, *96*(3), 248–260.

Pfurtscheller, G., & Lopes da Silva, F. H. (1999). Event-related EEG/MEG synchronization and desynchronization: basic principles. *Clinical neurophysiology*, *110*(11), 1842–1857.

Phillips, C. (2000). *Source estimation in EEG*. Unpublished doctoral dissertation, University de Liege, Belgium.

Plonsey, R., & Heppner, D. B. (1967). Considerations of quasi-stationarity in electrophysiological systems. *The Bulletin of mathematical biophysics*, *29*(4), 657–664.

Powell, M. J. D. (1969). A method for nonlinear constraints in minimization problems. In R. Fletcher (Ed.), *Optimization* (pp. 283–298). Academic Press.

Recht, B., Fazel, M., & Parrilo, P. A. (2010). Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization. *SIAM review*, *52*(3), 471–501.

Rockafellar, R. (1997). *Convex analysis*. Princeton University Press.

Rugg, M. D., & Coles, M. G. (1995). *Electrophysiology of mind: Event-related brain potentials and cognition.* Oxford University Press.

Sanei, S., & Chambers, J. A. (2008). *EEG signal processing*. John Wiley & Sons.

Sarvas, J. (1987). Basic mathematical and electromagnetic concepts of the biomagnetic inverse problem. *Physics in Medicine and Biology*, *32*, 11–22.

Schmidt, M., Roux, N. L., & Bach, F. R. (2011). Convergence rates of inexact proximal-gradient methods for convex optimization. *Advances in neural information processing systems*, *24*, 1458–1466.

Sharbrough, F., Chatrian, G.-E., Lesser, R., Lüders, H., Nuwer, M., & Picton, T. (1991). American Electroencephalographic Society Guidelines for Standard Electrode Position Nomenclature. *Journal of Clinical Neurophysiology*, *8*, 200–202.

Sörnmo, L., & Laguna, P. (2005). *Bioelectrical signal processing in cardiac and neurological applications*. Academic Press.

Sprechmann, P., Ramirez, I., Sapiro, G., & Eldar, Y. C. (2011). C-hilasso: A collaborative hierarchical sparse modeling framework. *Signal Processing, IEEE Transactions on*, *59*(9), 4183–4198.

Sra, S., Nowozin, S., & Wright, S. (2011). *Optimization for Machine Learning*. MIT Press.

Sriperumbudur, B., & Lanckriet, G. (2009). On the convergence of the concave-convex procedure. *Advances in neural information processing systems*, *22*, 1759–1767.

Stahlhut, C. (2011). *Functional Brain Imaging by EEG: A Window to the Human Mind*. Unpublished doctoral dissertation, Technical University of Denmark, Department of Informatics and Mathematical Modelling, Denamrk.

Stopczynski, A., Larsen, J., Stahlhut, C., Petersen, M., & Hansen, L. (2011). A smartphone interface for a wireless EEG headset with real-time 3D reconstruction. *Affective Computing and Intelligent Interaction*, 317–318.

Stopczynski, A., Stahlhut, C., Larsen, J. E., Petersen, M. K., & Hansen, L. K. (2014). The smartphone brain scanner: A portable real-time neuroimaging system. *PloS one*, *9*(2), e86733.

Stopczynski, A., Stahlhut, C., Petersen, M. K., Larsen, J. E., Jensen, C. F., Ivanova, M. G., ... Hansen, L. K. (2014). Smartphones as pocketable labs: Visions for mobile brain imaging and neurofeedback. *International Journal of Psychophysiology*, *91*(1), 54–66.

Sur, S., & Sinha, V. (2009). Event-related potential: An overview. *Industrial psychiatry journal*, *18*(1), 70.

Taniguchi, M., Kato, A., Fujita, N., Hirata, M., Tanaka, H., Kihara, T., ... others (2000). Movement-related desynchronization of the cerebral cortex studied with spatially filtered magnetoencephalography. *Neuroimage*, *12*(3), 298–306.

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society. Series B (Methodological)*, 267–288.

Trujillo-Barreto, N. J., Aubert-Vázquez, E., & Valdés-Sosa, P. A. (2004). Bayesian model averaging in EEG/MEG imaging. *NeuroImage*, *21*(4), 1300–1319.

Turlach, B. A., Venables, W. N., & Wright, S. J. (2005). Simultaneous variable selection. *Technometrics*, *47*(3), 349–363.

Turovets, S., Volkov, V., Zherdetsky, A., Prakonina, A., & Malony, A. D. (2014). A 3D Finite-Difference BiCG Iterative Solver with the Fourier-Jacobi Preconditioner for the Anisotropic EIT/EEG forward problem. *Computational and Mathematical Methods in Medicine*, *2014*.

Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., ... Joliot, M. (2002). Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage*, *15*(1), 273–289.

Uutela, K., Hämäläinen, M., & Somersalo, E. (1999). Visualization of magnetoencephalographic data using minimum current estimates. *NeuroImage*, *10*(2), 173–180.

Wing, G. M. (1991). *A primer on integral equations of the first kind: the problem of deconvolution and unfolding*. SIAM.

Wolters, C., Anwander, A., Tricoche, X., Weinstein, D., Koch, M., & MacLeod, R. (2006). Influence of tissue conductivity anisotropy on EEG/MEG field and return current computation in a realistic head model: a simulation and visualization study using high-resolution finite element modeling. *NeuroImage*, *30*(3), 813–826.

Wolters, C., Kuhn, M., Anwander, A., & Reitzinger, S. (2002). A parallel algebraic multigrid solver for finite element method based source localization in the human brain. *Computing and visualization in science*, *5*(3), 165–177.

Wolters, C. H., Grasedyck, L., & Hackbusch, W. (2004). Efficient computation of lead field bases and influence matrix for the FEM-based EEG and MEG inverse problem. *Inverse problems*, *20*(4), 1099.

Wright, S., & Nocedal, J. (1999). *Numerical optimization* (Vol. 2). Springer New York.

Wright, S. J., Nowak, R. D., & Figueiredo, M. A. (2009). Sparse reconstruction by separable approximation. *Signal Processing, IEEE Transactions on*, *57*(7), 2479–2493.

Yaghmour, K. (2013). *Embedded Android: Porting, Extending, and Customizing*. " O'Reilly Media, Inc.".

Yuan, M., & Lin, Y. (2006). Model Selection and Estimation in Regression with Grouped Variables. *Journal of the Royal Statistical Society. Series B, Statistical Methodology*, 49–67.

Zangwill, W. (1969). *Nonlinear programming: a unified approach*. Prentice-Hall.

Zhao, Y.-B. (2012). An approximation theory of matrix rank minimization and its application to quadratic equations. *Linear Algebra and its Applications*, *437*(1), 77–93.