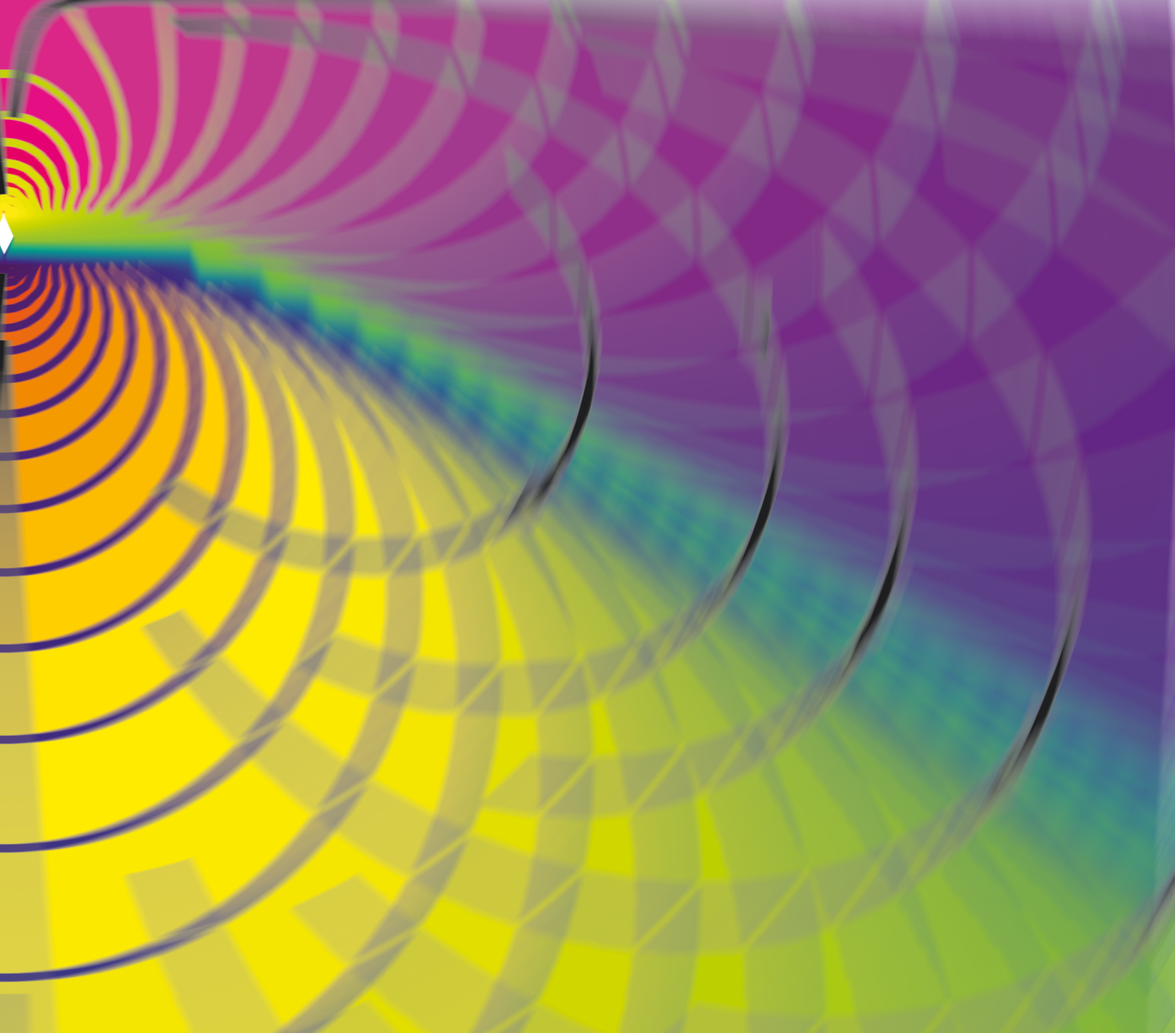

The Neurophysiology of EEG and the Physics of the Head

Theory and Application for Spontaneous EEG

Doctoral Thesis

Daniel Miklody

Technische Universität Berlin

The Neurophysiology of EEG and the Physics of the Head

Theory and Application for Spontaneous EEG

vorgelegt von
M.Sc.
Daniel Miklody
ORCID: 0000-0002-8709-4200

an der Fakultät IV - Elektrotechnik und Informatik
der Technischen Universität Berlin
zur Erlangung des akademischen Grades

Doktor der Naturwissenschaften
- Dr. rer. nat. -
genehmigte Dissertation

Promotionsausschuss:

Vorsitzender: Prof. Dr. Henning Sprekeler
Gutachter: Prof. Dr. Benjamin Blankertz
Gutachter: Prof. Dr. Gabriel Curio
Gutachter: Dr. Guido Nolte

Tag der wissenschaftlichen Aussprache: 16. Juni 2020

Berlin 2020

Abstract

Neuroscientific research using Electroencephalography is one of the most important tools for understanding human brain function and dysfunction. Not many other methods can non-invasively and directly access neural activity with millisecond precision. While basic event related modulations of brain activity can easily be replicated, the spatial resolution remains poor. Thus, complex or higher brain function and detailed aspects are much more challenging to analyze. The properties of volume conduction in the human head drastically decrease the spatial resolution of EEG. The mixing of neuronal sources on the scalp is linear but spatially smeared. EEG is highly autocorrelated while additionally non-stationary and non-linear in its emergence. Special methods are needed to differentiate sources based on temporal, spectral and spatial considerations. Many approaches based on common assumptions fail in practice. We need new ideas and a paradigm shift towards new perspectives in order to advance technology. This thesis introduces a new theory for the interpretation and classification of neural signals and develops algorithms based on it. The theory includes new perspectives on volume conduction as well as the propagation of oscillations and resonances in the cortex. This thesis suggests novel approaches to models of volume conduction, spatial filtering and optimal classifiers. The new perspective on volume conduction is based on but not limited to impedance measurements and includes approximate head model, sensor position and homogeneous conductivity estimations. In spatial filtering, novel side constraints on the common spatial patterns algorithm are investigated. Optimal Bayesian classifiers are derived for the direct classification of variance data and are related to established approaches based on the logarithm of the variances or Riemannian Geometry. The new perspectives on oscillations and resonance can be linked to the genesis of spectral harmonics due to the non-linear relation of synaptic input and firing frequency in single neurons. The results show the need for new approaches in head modeling, the interpretation of oscillations, spatial filtering and classification. Furthermore, they deliver implications for the investigation of functional connectivity and neuronal dynamics: the brain is a large musical instrument with finely tuned resonances in various spectral and spatial modes where the current harmony is based on the present and past perceptual input. The single resonances non-linearly depend on each other which leads to the necessary emergence of harmonics based on the principles of compression and soft-clipping.

Kurzbeschreibung

Neurowissenschaft mit Hilfe des Elektroenzephalogramm (EEG) ist eines der wichtigsten Werkzeuge zum Verständnis menschlicher Hirnfunktion und -dysfunktion. Nur wenige andere Methoden können nicht-invasiv und mit der Präzision von Millisekunden direkt auf die neuronale Aktivität zugreifen. Während einfache ereignisbasierte Veränderungen von Hirnaktivität leicht replizierbar sind, ist die räumliche Auflösung des EEG nicht optimal. Die Analyse detaillierter Zusammenhänge sowie komplexer und höherer Hirnfunktion bleibt eine Herausforderung. Die Volumenleitung führt zu einer schlechten räumlichen Auflösung des EEG und hat die starke Vermischung neuronaler Quellen zur Folge, während diese jedoch glücklicherweise linear ist. EEG weist durch dies und die spezielle Form seiner Nicht-Stationarität eine hochgradige Autokorrelation auf, ist jedoch zusätzlich nicht-linear in seiner Entstehung. Spezielle Forschungsansätze sind notwendig um die unterschiedlichen Quellen unter zeitlichen, räumlichen und spektralen Gesichtspunkten zu trennen. Da viele herkömmliche Ansätze versagen, braucht es neue Ideen und Perspektiven um die EEG-Technologie weiterzuentwickeln. Diese Arbeit stellt neue Theorien zur Interpretation und Klassifikation neuronaler Signale und darauf basierende Algorithmen vor. Die Theorien umfassen neue Aspekte der Volumenleitung sowie zur Entstehung und Weiterleitung von Oszillation und Resonanzen im Kortex. Biophysikalische Kopfmodelle, räumliche Filter und optimale Klassifikatoren werden daraus abgeleitet. Die Ansätze zur Volumenleitung basieren auf Impedanzmessungen, bieten jedoch darüber hinaus Näherungslösungen für individuelle Kopfmodelle sowie die Schätzung von Sensorpositionen und homogener Gewebeleitfähigkeiten. Neuartige Randbedingungen in der räumlichen Filterung werden untersucht. Bayessche optimale Klassifikatoren für Varianzdaten werden hergeleitet und in Bezug zu etablierten Ansätzen gesetzt, die auf dem Logarithmus der Varianzen oder Riemannscher Geometrie basieren. Die Entstehung im EEG messbarer spektraler Oberwellen wird auf Grund herkömmlicher Annahmen über den Zusammenhang von Feuerrate und synaptischem Input einzelner Neuronen aufgezeigt. Diese Arbeit zeigt die Notwendigkeit neuer Perspektiven in der Modellierung der Volumenleitung, der Interpretation von Oszillationen, der räumlichen Filterung und der Klassifikation. Die Ergebnisse eröffnen Implikationen für die Untersuchung funktionaler Konnektivität und neuronaler Dynamik: Das Gehirn ist ein riesiges musikalisches Instrument mit fein abgestimmten Resonanzen in vielfältigen räumlichen und spektralen Moden, dessen momentane Harmonie auf gegenwärtigem und vergangenem perzeptuellem Input basiert. Die Resonanzen hängen nicht-linear voneinander ab, was zur notwendigen Entstehung von Oberwellen durch die Prinzipien der Kompression führt.

Preface

Writing a dissertation is a long winding road that does not start at formulating the actual first word. The writing itself is the intensely focused climax of gathering information, discussing and formulating theory, generating insights through others as well as yourself. But all of this is impossible without education from many different sources. While the time of writing is often dominated by social isolation - intensified by the Corona virus in this case - even then it is these inputs and discussions that are the most important review of your thoughts and work.

If we believe that what we know about what people call 'reality' is roughly true, we are still enslaved by our senses to experience a world we actually cannot even prove to exist. All we are left with is a conscious perception of a surrounding that seems heavily influenced by our own experiences and expectations but nevertheless by external, by environmental factors. We perform actions to examine and manipulate this environment, investigating the relationships and the 'truth' behind our perceptions. This is inline with studying the brain "from inside out" as György Buzsáki terms it [Buzsáki, 2019]. Gathering rules and relationships that explain and sort the chaos around us seems inherent to all cultures. Science, however, is not free from beliefs and axioms that have to be trusted while simultaneously questioned as a basis for meaningful results.

Laplace, Lagrange and Legrendre first rejected Fourier transform, finding it mathematically arbitrary to describe functions by a functional transform with sinusoids [Zayed, 2019]. However, Fourier transform is now the most widely used spectral decomposition. This can be credited at least partly to the fact that our brain represents perceptions in a similar, spectral way and, hence, the transform is a good estimator for our expectations and perceptions.

Among many things in nature, our brain is simply a network of oscillations playing harmonics like a big improvising free jazz orchestra. Or maybe that accounts for the authors current brain state only. Brains might be playing classical music with a strict and relentless conductor or simply chaotic disharmonics depending on the person, phase in life, the day and the time of the day.

Nothing is universally true, it is shaped by our phantasy and this does not stop at the author's theories. You can only combine and formulate what you learn, see and hear as well as discuss it with others to not get trapped within your own small world.

Acknowledgments

Commonly, family is the initial funding that leads us into life, equips us with basic attitudes and self-awareness that form the most important parts of our character. So, first of all, I want to express my gratitude towards my parents and brother for providing me with such an open-minded, diverse childhood and youth that was heavily enriched by emotional warmth, humanistic attitudes, cultural openness as well as patient explanations and understanding.

The highest level of gratitude concerning this thesis I feel for you, Sarah, for your support, patience and understanding during the last months and years. I am incredibly happy that we have become what we are now. You are the basis of what I associate with home and family.

On a completely different level are the prospects that were enabled by my supervisor, Prof. Dr. Benjamin Blankertz. He always knows how to support me and discuss my research and ideas while introducing me to science and teaching: you have an incomparable gift of encouraging and supporting people. I want to thank Prof. Dr. Lucas C. Parra and the Neural Engineering Group at the City College of New York for hosting me and supporting me in developing what is now the linear leadfields approach. I owe a great thank you to Prof. Dr. Klaus-Robert Müller, Prof. Dr. Klaus Gramann, Prof. Dr. Gabriel Curio, Dr. Stefan Haufe, Dr. Andreas Ziehe and Dr. Guido Nolte: thank you for teaching, fruitful discussions and input during the last years.

Among the many great people in the environment of the Machine Learning Group and the Neurotechnology Group of TU Berlin that I call my friends and colleagues, I thank in particular Alex, Steffi, Markus, Max, Nils, Vera, Oleksandr, Sven, Johannes, MSK, Hendrik, Duncan, Dominik, Nadine, Miriam, Imke, Milena for the scientific input, discussions, help and a great time. I want to thank all the students that took part in my courses and were highly open for discussion.

I also want to thank nature. It is an important factor for me to always come back to the essences in life, to experience and explore nature - the basis of our complex being.

Contributions

The general contributions of this thesis are to be found in the development of new theoretical perspectives and the application of these insights into new methodologies. The description of the state-of-the-art is based on a vast inter-disciplinary literature review but not limited to the replication of contents. It enriches this literature review with own simulations and new theories that shed light onto commonly overseen aspects. Incorporating these insights into new approaches and theories leads to the development of novel classification approaches for variance data, spatial filtering techniques for EEG and head models with reduced complexity in the model and its creation.

Detailed distinction of significantly original contributions :

- Section 2.1: Novel simulations of volume conduction involving the method of images and a self-developed 4-shell BEM head model
- Section 2.2: New perspectives on the relation between a point processes and spectra. Theory on the emergence of harmonics based on the non-linear dependency of firing rate on the synaptic input
- Section 3.2.1: Novel approaches to realistic head model extraction [2]
- Section 3.3.2: Introduction of generalized χ^2-distributions, their application to variance data and new insights in LDA on the logarithm of variances
- Section 4.1: Development and application of a novel spatial filtering approach based on Common Spatial Patterns
- Section 4.2: Development and application of new optimal Bayesian classifiers for variance data
- Section 5.1: Introduction of Newtonian optimization to the estimation of conductivity in Boundary Element Method head models
- Section 5.2: Estimation of EEG electrode position based on impedances [1]
- Section 5.3: Novel dimensionality reduction for geometric head models [3]
- Section 5.4: A linear approach to estimating the individual head model based on impedance measurements

Included Contributions

[1] Miklody, D., Bagdasarian, M. T., and Blankertz, B. (2017a). A low-dimensional representation for individual head geometries. In *2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 3696–3699. IEEE.

[2] Miklody, D., Huang, Y., Haufe, S., and Parra, L. (2016a). Automatic creation of fast, realistic boundary element head models. In *2016 22nd Annual Meeting of the Organization for Human Brain Mapping (OHBM)*.

[3] Miklody, D. and Höhne, J. (2015). Impedance based automatic electrode positioning. In *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 646–649.

Additional Contributions

(i) Miklody, D., Moessmer, P., Dettmann, T., Klinkenberg, K., and Blankertz, B. (2017b). Multi-timescale spectra as features for continuous workload estimation in realistic settings. In *Proceedings of The 7th International BCI Conference Graz, 2017*.

(ii) Miklody, D., Uitterhoeve, W. M., Heel, D. v., Klinkenberg, K., and Blankertz, B. (2016b). Cognitive workload bci in the maritime environment. In *Proceedings Of The 6Th International Bci Meeting: Bci Past, Present, And Future*.

(iii) Miklody, D., Uitterhoeve, W. M., van Heel, D., Klinkenberg, K., and Blankertz, B. (2016c). Maritime cognitive workload assessment. In *International Workshop on Symbiotic Interaction*, pages 102–114. Springer.

(iv) Sturm, I., Treder, M., Miklody, D., Purwins, H., Dähne, S., Blankertz, B., and Curio, G. (2015). Extracting the neural representation of tone onsets for separate voices of ensemble music using multivariate EEG analysis. *Psychomusicology: Music, Mind, and Brain*, 25(4):366.

(v) Treder, M. S., Purwins, H., Miklody, D., Sturm, I., and Blankertz, B. (2014). Decoding auditory attention to instruments in polyphonic music using single-trial eeg classification. *Journal of neural engineering*, 11(2):026009.

Contents

Chapter 1

Introduction

In neuroscience, great achievements have been made during the last century since the first recorded Electroencephalography (EEG) by Berger in the 1920s. While many of the fundamental theoretical concepts date back to even earlier and many foundations of theoretical neuroscience have also been laid during the mid-century, most technical advances date later or around the turn of the millennium.

Magnetoencephalography (MEG) was invented around 1968 and has helped gain great insights into the function of the neocortex. Functional Magnetic Resonance Imaging (fMRI) was only developed around 1990, opening a whole new window on brain function and connectivity. Diffusion Weighted Imaging (DWI) as a technique also based on MRI, developed around 1985, found its way into brain research approximately in the year 2000. It is now used to estimate conductivities, the functional connectivity and myelinization within the brain.

The increasing computational power and the variety of approaches drastically changed the analysis since the 1980s. In particular the evolution of machine learning culminated around 2010 in the expectancy of being able to control almost anything with brain computer interfaces (BCIs) within a few years and to build artificial brains.

Since then, openly speaking, not too much has happened that really advanced the field of BCI. The main challenges remain the same while new approaches such as applying deep learning on EEG data or the development of classifiers based on Riemannian Geometry can improve performance, but, sadly, not substantially.

The novelties in theories in the last decade are mainly related to the development of new paradigms explaining in particular brain connectivity and dynamics. The evolution of computational neuroscience and the further increase particularly in parallel computing power opens new fields. But, also, new perspectives on oscillations and background activity of the brain among the insight that the thickness of myelinization within in the brain is depending on learning effects lead to a manifestation of the importance of brain dynamics.

This thesis is born out of necessity for integrating theoretical considerations and advances into the applied methods in order to build valuable new approaches. The main focus lies in the investigation of and application to spontaneous EEG, while many of the results also apply to event-related potentials (ERPs) and similar methods.

In *Chapter 2*, we start from neurons - the main electric field generators within the brain - and investigate the resulting external field based on volume conduction. We then take a detour via the statistical implications of point processes like action potentials and related synaptic potentials, in order to explain the emergence of certain spectral phenomena that we observe in EEG, MEG and intracranial recordings. As a result of these theoretical considerations, a foundation for the explanation of spectral harmonics is laid, but, also other interpretations are revisited in a common theory.

Chapter 3 first introduces methods to measure neural activity and estimate the individual electrical properties of the human head. Details on standard analysis of spontaneous EEG data are reviewed. The statistics of variances are investigated which results in a novel perspective on one of the most common approaches, Linear Discriminant Analysis (LDA) on the logarithm of variance, and its relation to Riemannian geometries.

In *Chapter 4*, a novel spatial filtering technique is introduced and evaluated on motor imagery BCI data, after which a new family of classifiers is derived and evaluated on the same data set. The classifiers are optimal Bayesian classifiers for χ^2 distributions and related linear approximations.

Chapter 5 deals with novel approaches to approximating tissue conductivities, sensor positions, geometric headmodels and EEG leadfields. Also, insights on technical details are shared, that regard the essence of good correspondence in the electrode, source and head geometry estimations for a proper estimation by approximate methods.

The key contributions of this thesis are novel theories of head model approximations via impedance measurements, new approaches to the investigation of oscillations, novel classifiers and spatial filters among a generic framework for geometric dimensionality reduction.

Chapter 2

Theory of Electrical Signal Generation of Neurons

Electroencephalography (EEG) measures the electrical potential on the scalp in order to gain information with millisecond precision directly linked to the electrical activity caused by propagation of impulses from neuron to neuron. While it is under strong influence of artefacts and thus suffers from a low Signal-To-Noise ratio (SNR), we will focus here on the neural origins of EEG and treat the latter in the next chapter.

Commonly, the question arises which part of neural activity are measured and from which areas of the brain. Presumably, predominantly assemblies of neocortical pyramidal cells cause signals spatially large scale and strong enough to be visible in EEG. Of their various small scale signals, the post-synaptic potentials (PSPs) are primarily measured, but also hyper-polarization and action potentials (APs) can be measured under certain circumstances.

In general, many more components take part in the neural processing of the brain. These components are mainly other types of neurons - like stellate cells and interneurons - as well as Glia cells but also involve subneuronal levels of activity. Within this chapter we will start from the fields generated by single neurons, advance to their temporal and thus spectral structure in elicited external electric fields and then discuss the effects of distributed sources including their functional coupling.

2.1 Neurophysiology & Volume Conduction

Let us start with a single neuron and investigate its role in the electrical and the neuronal signal propagation. Neurons are thought to be the main local node in the neural processing taking place in the brain. There are, however, sub-neuronal levels of investigations down to single molecules or below. Processes on this level contribute to or modulate neural activity but the EEG in particular is only sensitive to the activity spread over a large number of neurons. This leads to the focus on an investigational level from neurons upscale within this thesis. Also, we will focus on the main contributors to EEG activity while sources with minor impact and other recording methods will be mentioned.

2.1.1 Sources and Models of Neuroelectric Activity

If there exists an electrical field measurable as potential in a conductive medium like biological tissue, electrical current has to flow. Below 1 kHz, the investigations can be reduced to quasi-electrostatic approximations [Häamäläinen et al., 1993]. This implies that a separation of charges leads to a current flow towards equalizing the so produced potential difference. In the brain, the main conductive elements are the small charged ions such as sodium (Na^+), potassium (K^+), chlorine (Cl^-) and calcium (Ca^{2+}). These are also involved in the signaling between neuronal cells and muscle activation among many other general purposes in the human body. The membranes of neurons are electrically charged by active and passive transport mechanisms in shape of ion channels that lie within the cell membrane. Neurons have a negative resting potential of around -65 mV that is in principle maintained by active pumps producing a concentration gradient of Na^+ and K^+.

When neurons receive synaptic input, the membrane potential changes which in turn triggers active and passive mechanisms in the membrane that lead to an in- and outflux of differently charged ions along the membrane. This triggers secondary currents restoring electroneutrality in the extra- and intracellular space which are focused mainly along the dendrites for PSPs and along the axon for APs. These external return currents of the primary currents through the membrane are what is commonly termed volume conduction due to its distribution over the whole surrounding conducting media [Buzsáki et al., 2012].

2.1.1.1 Neuronal Sources of Electrical Fields

In *Figure 2.1*, we see a sketch of a pyramidal cell with five main sources and connected current loops that resemble the main components of the neural activity linked to their electrical far field. The sketch is a strong simplification and the sources and sinks are only to be seen exemplarily. Among these current sources, there are two groups: synaptic input on the dendrites and the action potential on the axon. The external currents differ in spatial spread and orientation, synchronicity and time constants.

The main neuronal sources of EEG potentials are commonly related to spatially spread *excitatory and inhibitory post-synaptic potentials* (EPSPs/IPSPs) producing an extracellular return current along the dendrites [Buzsáki et al., 2012]. Notably, the probability of inhibitory vs. excitatory synapses is location dependent. In CA1 pyramidal cells in the Hippocampus, the distal apical dendrites have a ratio of 86% to 83% of excitatory vs inhibitory synapses with a density up to $1.72m^{-1}$. While advancing to the soma on the thick distal radiatum first increases the ratio up to 98% and the density to $6.9m^{-1}$, it thereafter decreases down to 2% and a density of $1.7m^{-1}$ approaching the soma [Megıas et al., 2001].

The propagation of PSPs along the dendrites is mainly passive but their influence is

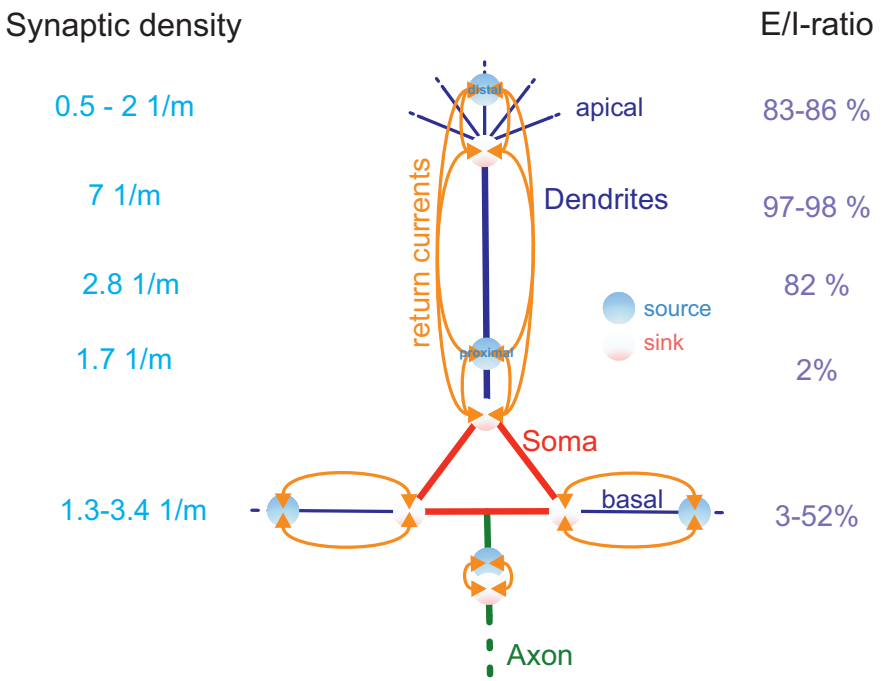

Figure 2.1: External return currents of a hippocampal CA1 pyramidal cell: the main far-field produced by a pyramidal cell comes from post-synaptic potentials in the apical dendrites as all the other fields are on average closed by random orientation (basal dendrites) or very local (action potential on the axon). Synaptic density and excitation/inhibition ratio (E/I) are adopted from [Megıas et al., 2001]

linearized by certain active mechanisms [Cash and Yuste, 1999]. In the axons, however, action potentials, as an exclusive all-or-none event, dominate, leading to the limited spatial spread of PSPs over dendrites and somata. The distribution of excitatory and inhibitory synapses in distal and proximal areas to the soma (see *Figure 2.1*) leads to a similar external field of both due to the oppositely charged currents. Inhibitory synapses are mostly located close to the soma, while excitatory in the distal parts of the apical dendrites. The propagation along the dendrites is thus in opposite directions but due to the opposite polarity, EPSPs and IPSPs have a similar external field.

Additionally, the *action potentials (APs)* of spike bursts can have an effect on higher frequencies of EEG e.g. in the somatosensory evoked potentials (SEPs) [Curio et al., 1994; Baker et al., 2003; Gobbelé et al., 1998] or in epileptic seizures [Frauscher et al., 2017]. The fact that they are additionally very local makes them challenging to measure in EEG as their low amplitude in combination with spatial smearing mostly disappears in particular in the noise of the amplifiers [Scheer et al., 2005]. Electrocorticography

(ECoG) and local field potentials (LFP) reveal further information on these more local phenomena.

Another main source of EEG activity is what is closely related to the so-called *Bereitschaftspotential*. It is commonly concluded that the reason for this is not a direct post-synaptic potential but an afterhyperpolarization of the cell membrane following excessive calcium-induced firing (bursts) [Hotson and Prince, 1980; Buzsáki et al., 2012]. The resulting passive currents along the dendrites produce similar field configurations as those of the PSPs, which in turn leads to the usage of similar dipole models [Böcker et al., 1994; Murakami and Okada, 2006].

In general, many effects of neural activity produce a local electric field (often termed equivalently to the measurement method, the local field potential - LFP), but they are usually not strong enough to be measurable by the far field in EEG and thus need more local recording techniques to be identifiable.

For most of these sources, the equivalent current dipole (ECD) model is used as first approximation, which we will take a closer look at in the next section.

2.1.1.2 The Equivalent Current Dipole

The different sources are commonly modeled by the so-called equivalent current dipole (ECD) for EEG as well as Magnetoencephalography (MEG). The ECD approximates the synchronized activities for sub-patches of cortex with common orientation of sources. The equations for a dipolar source in general can be set up equivalently for the electrical scalar potential ϕ_{dip} and the source current density $\vec{J}_{dip}$:

$$\phi_{dip}(\vec{r}) = \frac{1}{4\pi\epsilon}\frac{\vec{p}\cdot\vec{r}}{|\vec{r}|^3} \tag{2.1}$$

$$\vec{J}_{dip}(\vec{r}) = -\sigma\nabla\phi = \frac{\sigma}{4\pi\epsilon}\left(3\frac{\vec{p}\cdot\vec{r}}{|r|^5}\cdot\vec{r} - \frac{\vec{p}}{|r|^3}\right) \tag{2.2}$$

Given here are the analytical solutions for an infinite homogeneous medium with the source at the origin that are usually the basis of more complex models incorporating different levels of realism of the surrounding media. $\vec{r}$ is the position relative to the dipole and $\vec{p} = |p|\,\vec{e}_p$ is the dipolar moment of the sub-patch of cortex, consisting of the dipole strength $|p|$ and the mean field orientation $\vec{e}_p$, which can alternatively be expressed in a similar way directly related to the local currents [Murakami and Okada, 2006]. Conductivity is $\sigma = 0.25S/m$ and the relative permittivity is $\epsilon_r = 10^6$ for the neuronal source space (the gray matter of the neocortex) [Gabriel et al., 1996; Miklavcic et al., 2006], where the absolute permittivity is $\epsilon = \epsilon_r\epsilon_0$.

The larger the spatial spread $\vec{l}$, the larger is the amplitude in the far field. Hence, the action potential with its more local structure has much less effect on the far field, in particular for EEG, as it is a very fast process on a small timescale rarely in synchrony

with similarly oriented neighbors. The current produced by a single neuron is too small to be measurable in the far field. A single cortical pyramidal cell produces currents in the range of 0.2pA [Murakami and Okada, 2006] while an equivalent current dipole as estimated by MEG [Häamäläinen et al., 1993] is in the range of 10nA. Hence, the summation of large numbers (10 000 to 50 000, see also [Murakami and Okada, 2006]) of synchronously active and similarly oriented neurons lead to a dipolar structure in the far field of those neurons strong enough to be measured by EEG or MEG.

It must be said, however, that the ECD model is a helpful approximation, as we have a net-zero source and sink structure which coarsely resembles this dipole. However, it is a valid in the far field only due to the distributed return currents along the dendrites. In particular in the near-field of LFPs, relations are far more complex. Here, multi-pole models or models of local ionic membrane currents at the synapse and passive cable equations seem to be the minimum to receive a satisfactory field description [Hines and Carnevale, 1997].

2.1.2 Volume Conduction

2.1.2.1 Phenomenological Observations

In order to qualitatively investigate the basic effects of volume conduction, we will start off with a simple dipole in an infinite homogeneous medium. In *Figure 2.2*A, we can see the distribution of equipotential lines and the connected vector field for a dipole in the center. In general, the head consists of a mixture of different tissues and not an infinite homogeneous medium. The resulting field in *Figure 2.2*B and C illustrates the local situation of a pyramidal cell close to the cortex surface. This is a very simple example of what happens for a sudden change in conductivity: if the conductivity decreases from CSF to skull or in a 3-shell model between brain and skull (B), the current flow is distorted away (repelled) from the boundary in search for the least energy consuming path. While for an increase from brain to CSF (C), it is attracted, because more current flows through the more conductive medium. B is by the way also an example of the scalp-air interface effect.

Actually, the different tissues such as white and gray matter, Cerebrospinal Fluid (CSF), skull and scalp are quite inhomogeneous themselves and there are several sudden changes between those tissue types with a change to non-conductive media at the scalp which is rather the opposite of an infinite homogeneous medium. Additionally, we find anisotropies in the conductivity in particular in white matter [Wolters et al., 2006; Güllmar et al., 2010; Vorwerk et al., 2014].

Many papers have shown that the conductivities of the single tissue types have significant influence on amplitude and topography of scalp potentials, in particular those of the skull and scalp (e.g. [Haueisen et al., 1999; Nissinen et al., 2011; Fernández-Corazza et al., 2017; Ferree et al., 2000; Goncalves et al., 2003; Clerc et al., 2005b; Malony et al.,

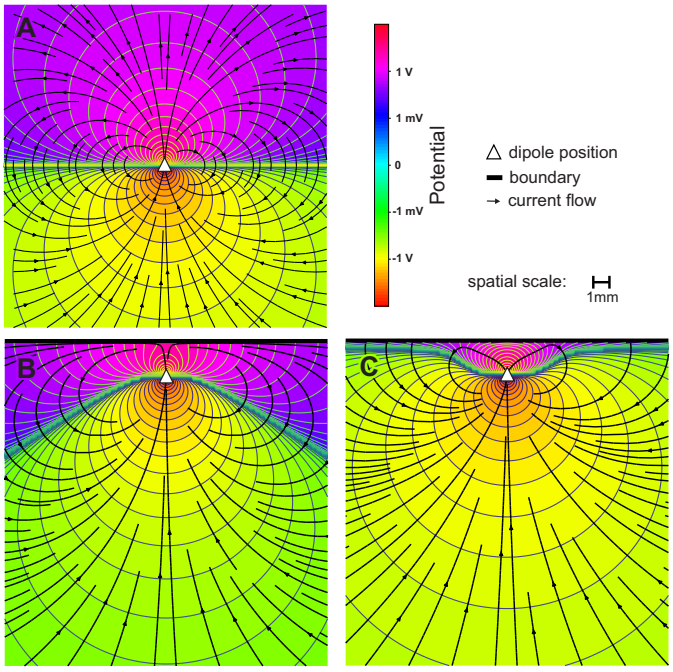

Figure 2.2: An upwards oriented current dipole ($I_0 = 1nA$) located at the white triangle in different media: **A** in infinite homogeneous medium ($\sigma_1 = 0.25S/m$ - brain tissue), **B** in semi-infinite homogeneous medium ($\sigma_1 = 0.25S/m$ - brain tissue) with a boundary to a less conductive medium ($\sigma_2 = 0.004125S/m$ - skull tissue): this resembles a pyramidal cell very close to the skull in a 3-shell head model, **C** in semi-infinite homogeneous medium ($\sigma_1 = 0.25S/m$ - brain tissue) with a jump to a much higher conductive medium ($\sigma_2 = 1.78S/m$ - CSF): this resembles a pyramidal cell close to the CSF in a 4-shell head model. The currents are repelled or attracted from a change in conductivity (black line).

2011; Datta et al., 2013]). This is essential for source localization and functional connectivity estimates as we will see in *Chapter 3* [Baillet et al., 2001; Fuchs et al., 2002; Darvas et al., 2006; Haufe et al., 2008; Valdés-Hernández et al., 2009].

Exemplarily, *Figure 2.3* shows one of the simplest approaches for realistic head models: Boundary Element Method (BEM) head models with different numbers of shells. The effects seen here are what most people refer to as the effects of volume conduction or spatial smearing. We can see that from adding the skull to the model (3-shell model), not only the amplitude decreases but also the focality and the clear dipolar structure are lost in the scalp potential. Further adding the CSF strengthens this effect (4-shell model). We can see that the main voltage is produced close to the source in the upper row and

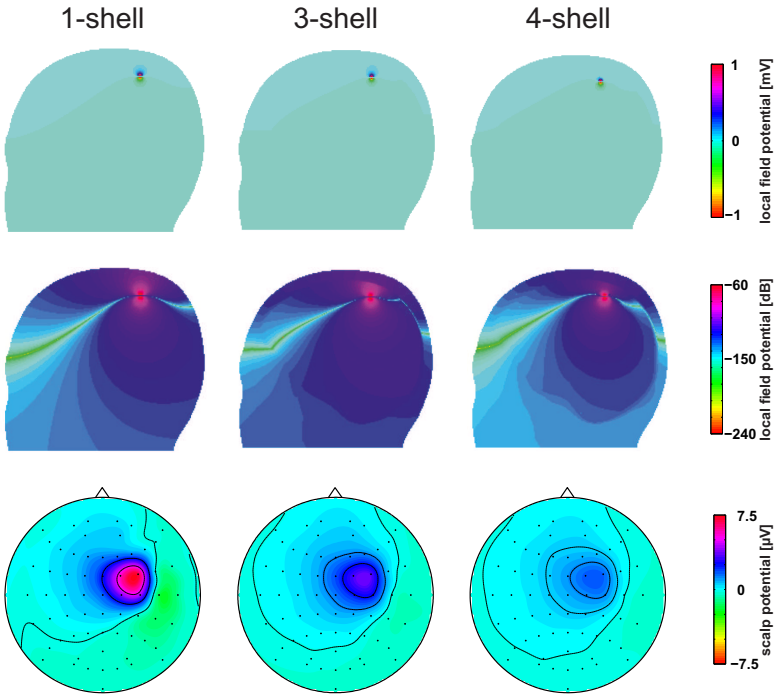

Figure 2.3: A current dipole in the primary motor area (M1) in Boundary Element Method (BEM) head models with different numbers of shells: shown is the potential in a sagital cut at the level of the source in lateral direction in the upper two and the scalp potential as measured by EEG on the lower row.

that the potentials are distorted in a similar way as in *Figure 2.2* by the boundaries between the tissue types (not shown).

Inspecting the potential distributions inside the head (central row), we find a deflection of the field away from the scalp surface in the 1-shell model which already distorts the zero potential line (practically the yellow center of the green area) downwards. As the exemplary dipole is perfectly oriented towards the superior direction (upwards), this line would be perfectly horizontal in an infinite medium. By this already, the scalp potential is slightly blurred and the dipolar structure is decreased. Adding the skull (3-shell model) increases this effect in particular inside the brain, the zero potential line is deflected further to inferior locations in this slice. In the 4-shell, we find two main effects: the zero-potential line inside the brain is suddenly more similar to that of the 1-shell but the field is further compressed due to the jump to a higher conductivity in the CSF as seen from the brain tissue. Inside of the skull, a lot of the current flows through the CSF due to its high conductivity which counteracts the repelling effects of skull and scalp surface (1 and 3-shell). However on the scalp surface the peak amplitude

is further lost and the dipolar structure further vanishes due to the shielding effect of the CSF: the equipotential lines are blurred even further.

These theoretical considerations underline the importance of modeling not only the skull but also the CSF within a reasonable head model. This need has also been found in numerical simulations [Ramon et al., 2004]. In general, also the inhomogeneity of the skull, the anisotropy of the white matter and further factors have to be considered for highly precise volume conductor models [Haueisen et al., 1997; Wolters et al., 2006].

2.1.2.2 Theoretical Background

The relation between currents and potentials in frequencies below 1kHz is commonly assumed to be constant and linear due to quasi-electrostatics. This resembles the local form of Ohms law and the possibility of applying the superposition principle: the resulting field of several simultaneously active sources is simply the sum of their effects.

For the theoretical derivation we will start from the local form of the Maxwell Equations

$$\begin{aligned}\nabla \cdot \vec{D} &= \rho \\ \nabla \cdot \vec{B} &= 0 \\ \nabla \times \vec{E} &= -\delta_t \vec{B} \\ \nabla \times \vec{H} &= \vec{J} + \delta_t \vec{D},\end{aligned}$$

wherein the charge retention $\nabla \cdot J = -\delta_t \rho$ is implicitly included in the last equation. Here, $\vec{D}$ is the electric flux, $\vec{B}$ the magnetic flux, $\vec{E}$ the electric field strength and $\vec{H}$ the magnetic field strength. In frequencies below 1kHz, we assume electrostatics, as mentioned earlier, and thus neglect electromagnetic induction ($\delta_t \vec{B} = 0$) and assume a static space charge ($\delta_t \rho = 0$). This results in $\nabla \cdot \vec{J} = 0$: a divergence free current is found, meaning that there are no current sources in the field itself, only those of active sources like the neuronal membrane activity or externally applied currents. Further, we get the possibility of describing the electric field with a simple potential Φ: $\vec{E} = -\nabla\Phi$ as its curl is $\nabla \times \vec{E} = 0$ (the field is conservative).

A macroscopic model for the current density inside the head is the combination of this field description including an ohmic current $\vec{J}_{vc} = \sigma\vec{E}$ with a primary source current $\vec{J}^P$ (e.g. neuronal activity): $\vec{J} = \sigma\vec{E} + \vec{J}^P$ [Clerc et al., 2005b]. In this fashion the causes for the currents (e.g. diffusion or an electric field) are only treated indirectly in $\vec{J}^P$ while the resulting field of these currents lead to the ohmic volume conduction $\vec{J}_{vc} = \sigma\vec{E}$. If we want to solve these differential equations, we need to find the homogeneous solution for $\vec{J}_{vc}$ that solves the field for the imprinted partial solution $\vec{J}^P$.

The macroscopic model in combination with the charge retention leads to $\nabla \cdot J = \nabla \cdot \left(\sigma E + J^P\right) = \nabla \cdot \left(-\sigma\nabla\Phi + J^P\right) = 0$. Reordered, we get a Poisson equation suitable

for the EEG case [Sarvas, 1987]:

$$\nabla \cdot (\sigma \nabla \Phi) = f = \nabla \cdot J^P. \tag{2.3}$$

where f is a function of space describing the current source density. If we know the distribution of sources and hence the current source density, we can solve the linear system to receive the resulting field.

For injected currents through the scalp as in Transcranial Electrical Stimulation (TES) or Electrical Impedance Tomography (EIT), we consider the contribution of the inner sources as negligible and thus $f = 0$ [Clerc et al., 2005b]. This leads to the following Laplace equation inside the head:

$$\nabla \cdot (\sigma \nabla \Phi) = 0. \tag{2.4}$$

The basic solution $G(\mathbf{r}, \mathbf{r}')$ to a Laplace operator solves $\nabla^2 G(\mathbf{r}, \mathbf{r}') = -\delta(|\mathbf{r} - \mathbf{r}'|)$, where $\nabla^2 = \nabla \cdot \nabla$ and $\delta(x)$ is the Dirac delta. This solution is not necessarily harmonic and does in general not fulfill the boundary conditions of subspaces.

The Green function is such a solution. It also vanishes at infinity which is in correspondence with the physical intuition that a static field has no effect far away from its sources [Kybic et al., 2005]. In 3 dimensions the Green function G is:

$$G(\mathbf{r}, \mathbf{r}') = \frac{1}{4\pi |\mathbf{r} - \mathbf{r}'|}. \tag{2.5}$$

In contrast, for the problem described by a Laplace equation $\nabla^2 \Phi = 0$, the solution is a harmonic function (by definition) and is fully determined by the values on the boundaries.

The solution to these equations is unique, hence it does not matter how it is constructed. To construct a solution for any divergence function f of the current source density in a subspace e.g., we can first calculate the free space solution for $\nabla \cdot J^P = f$ and then add a harmonic solution for *Equation* (2.4) to fulfill the boundary conditions.

For source free subspaces the solution is a simple harmonic function determined by the boundary conditions. Also sources in several subspaces are possible. An exemplary numerical solution to these equations is introduced in *Chapter 3.2.1.3*.

The simulations in *Figure 2.2* B+C were created in a similar fashion: we first simulated the field of a dipole in an infinite homogeneous medium and then added the same potential spatially mirrored at the boundary weighted by the conductivities in order to fulfill the boundary conditions. The normal current has to have no jump on the boundary (is continuous) and in this case (connected conductive space) the potential also has to be continuous. Hence, we find jumps [ ] in the normal electrical field strength $\delta_n \Phi$ which is the normal derivative of the potential Φ and proportional to the normal current

density J_n by the conductivity change:

$$[\Phi] = 0 \tag{2.6}$$

$$[J_n] = [\sigma \delta_n \Phi] = 0 \tag{2.7}$$

For given discretized geometries, *Equation* (2.3) can be explicitly calculated and transformed into a linear equation of the form:

$$\mathbf{A}\mathbf{x} = \mathbf{b} \tag{2.8}$$

where $\mathbf{A}$ incorporates the spatial dependencies as convolutions with the Green function and $\mathbf{b}$ depicts the inhomogeneities as boundary conditions while $\mathbf{x}$ are either related to the normal currents or the potentials or a combination of both in different spatial locations depending on the type of model.

2.2 Oscillations & Power Spectral Densities

Now that we know how single neurons contribute to the field measured by EEG and that we need temporally synchronous activity for the mean field to be strong enough, we want to investigate how this synchrony arises. There are many different perspectives on EEG data, of which the oldest form is to look at the signal in the temporal domain. In these, already the pioneer of EEG Hans Berger in 1929 [Berger, 1929] detected a certain spontaneous rhythmic activity with two frequencies between 8-11 Hz and 25-30 Hz which he termed alpha and beta waves later on [Berger, 1931]. 1931 already, he had observed an increase in beta and a decrease in alpha corresponding to the increase in 'mental work'. These early observations suggest the investigation in the spectral domain.

Various local and global oscillations influence the propagation of neuronal pulses. These oscillations occur on all scales from membrane potentials of single neurons to networks spanning the whole nervous systems. However, also stochastic resonances [Baker et al., 1999; Jones, 2016], the non-sinusoidal shape of neural firing [Cole and Voytek, 2019; Pfurtscheller et al., 1997] and time-delayed same-frequency oscillations [Schaworonkow and Nikulin, 2019] enter the average power-spectral densities as peaks. Hence, a spectral peak does not necessarily mean an (ongoing) oscillation at that frequency but could have various reasons. But let us first look at the properties of power-spectral densities of EEG.

Power Spectral Densities (PSDs) $S_{xx}(\omega)$ decompose a stochastic signal into different sinusoidal component in their power usually using a windowed short-time Fourier transformation using the Welch's method [Welch, 1967]. It characterizes the average (or expected) power density per frequency ω in short time windows of a signal $x(t)$ of length τ in the frequency domain [Rieke et al., 1999]. If we denote by $\hat{x}(\omega)$ the Fourier

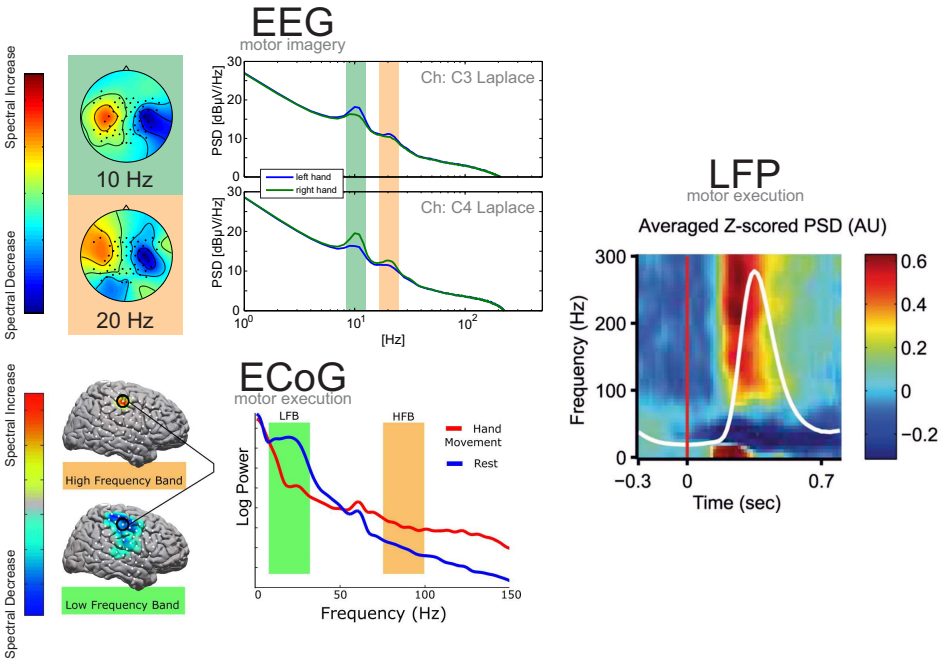

Figure 2.4: Typical spectra of EEG and ECoG and a spectrogram of LFP in hand movement tasks: we can observe a general shape of 1/f with certain oscillations peaking in the PSDs of ECoG and EEG. The peaks are not sharp in frequency suggesting a rather probabilistic oscillatory structure they are based on. The source patterns of the different frequency differences overlap while they suggest a decrease of the active areas. There is also a 3rd harmonic visible in EEG. The higher frequencies reveal more focal spatial patterns. The LFP changes mainly in higher (Gamma) frequencies while there is also a decrease in the beta range and an increase in lower frequencies in motor preparation (the white line is hand movement speed, the Z-score depicts the change from the mean standard deviation). Reprinted with permission: ECoG from [Miller et al., 2007], copyright 2007 Society for Neuroscience, and LFP from [Perel et al., 2015], copyright 2015 The American Physiological Society.

transform of $x(t)$ which is defined over a finite time interval T then:

$$S_{xx}(\omega) = \lim_{T\to\infty} \left\langle |\hat{x}(\omega)|^2 \right\rangle \tag{2.9}$$

which is equivalent to the Fourier transform $\hat{R}_{xx}(\omega)$ of the auto-correlation $R_{xx}(\tau)$ (Wiener-Kintchine-Theorem):

$$S_{xx}(\omega) = \hat{R}_{xx}(\omega) \tag{2.10}$$

In the PSD of EEG signals, we can observe:

- A general shape of $1/f^{\alpha}$ in the spectrum
- Peaks at different frequencies

Both peaks and the 'background noise' are found to be dependent on different cognitive states.

Oscillations can be investigated in the spontaneous EEG as classically done by Berger, but we can also investigate typical spectra in the event-related responses (ERP) connected to a repeated stimulus or similar event. The ERP in time domain consists of an oscillatory structure with different peaks, which seemingly have a basic frequency of around 10 Hz at least for the first peaks, while the later P300 peak is slower. The connection between spontaneous and event-related oscillations is highly under debate, but high correlations between alpha and ERPs can be found [van Dijk et al., 2010]. As the EEG is rather a macroscopic low-resolution signal of the brain - a complex non-linear dynamical system - the separation into single sources of these oscillations is a tough task.

There are several theories and studies on phase synchrony and/or reset between ERPs and spontaneous alpha oscillations [Risner et al., 2009; Milton and Pleydell-Pearce, 2016] that are highly debated [Ritter and Becker, 2009] while a connection is thoroughly found, e.g. between the patterns of the P300 and the alpha peak [Intriligator and Polich, 1994]. Also, the generation of ERD/ERS can be linked to ERPs [Yordanova et al., 2001]. We will focus on the investigation of spontaneous EEG activity throughout this thesis.

2.2.1 Single Event Sources of EEG in the Spectral Domain

We will now investigate how single events of neurons contribute to the field in the temporal and spectral domain. Although PSPs are the main event we measure in EEG, they are triggered by single events - the action potentials - and hence must obey the firing statistics of presynaptic neurons. As we will see, the point process like structure of action potentials leads to a frequency-independent (white) spectral structure by itself, while the type of triggered event (AP, IPSP or EPSP) determines the spectral shape together with the statistics of the input. Synapses are the communications channels between neurons and the brain is highly recurrent. So, if we measure certain oscillations they can only emerge from the functional connectivity between neurons in combination with the time delays of these feedback loops.

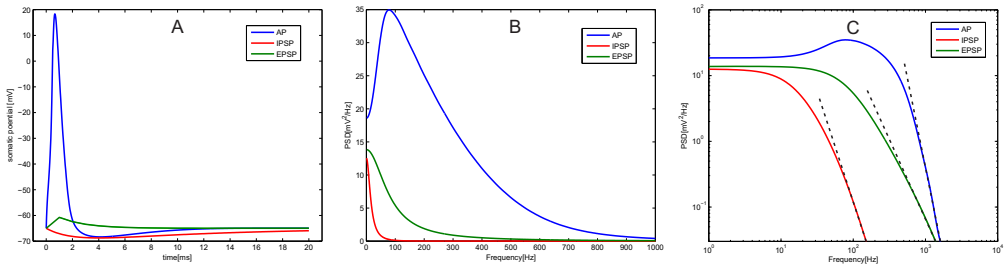

Figure 2.5: Typical time course (A), PSD (B) and double logarithmic PSD (C) of action potentials (AP) and postsynaptic potentials (PSP) simulated in a layer 5 pyramidal cell using NEURON [Hines and Carnevale, 1997]. The dashed lines depict the fractional exponent α in the high frequencies ($\alpha = 2.3$ EPSP, $\alpha = 2.7$ IPSP, $\alpha = 4.5$ AP). Note that the scaling was fitted for visibility.

2.2.1.1 Action potentials (AP)

Action potentials (AP) are the main bits of neuronal communication, transporting information within neurons along the axon. They also trigger the synaptic output from the different sensory cells as well as the output to the different muscles. Many studies suggest a non-deterministic stochastic firing of single neurons with a firing pattern in-vivo influenced by the different excitatory and inhibitory inputs as well as the current past of the neurons activity [Bair et al., 1994; Pesaran et al., 2002; Lampl and Yarom, 1993; Puil et al., 1994; Pike et al., 2000; Pedroarena and Llinás, 1997]. Firing statistics of single neurons are mostly characterized in inter-spike-interval (ISI) distributions which in-vivo seem to be related to a Poisson process on first approach [Dayan and Abbott, 2001]. Isolated single cells stimulated by current injections in-vitro, however, seem to be firing rather deterministically. This suggests that the reason for the irregularity rather lies in the large amount of random input from synaptic 'background' activity stochastically influencing the membrane potential [Holt et al., 1996; Buzsaki, 2006]. The resulting ISI distribution of spike times τ of such a Poisson process with a constant firing rate r is a simple exponential:

$$p_{ISI}(\tau) = re^{-r\tau} \tag{2.11}$$

Interestingly, if we simulate random (Poisson) synaptic bombardment as linear superpositions of many APs or PSPs and calculate the PSD of the resulting time signal, we will receive a fuzzy version of the original waveform (AP/PSP), see *Figure 2.6*. The reason is that if we take many action potentials in a random order as single events occurring repeatedly over time and calculate the PSD, the resulting signal is still linearly depending on the single events. The mathematical operation in time domain here represents a convolution with a delta pulses at different points in time which has a flat (white) frequency spectrum. Convolution in time domain represents a multiplication in

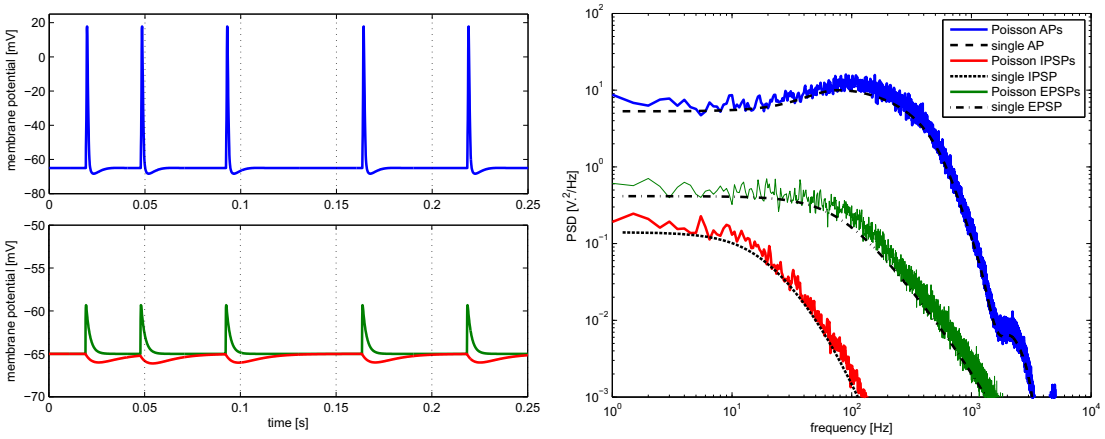

Figure 2.6: Timeline and PSDs of action potentials and corresponding EPSPs driven by a homogeneous Poisson process with 20Hz average firing rate. The Poisson process itself produces a white spectrum while the rest of the spectrum is determined by the shape of the single point process. Note that the scales of the PSDs depend on the rate and have been amplified for visualization.

frequency domain, hence we practically multiply it with a (in this imperfect case noisy) constant.

If the firing rate $r(t)$ varies over time, we find what we call an inhomogeneous Poisson process. It can be modeled similarly to a homogeneous Poisson process. The probability of firing a spike during a short interval of duration Δt is $r(t)\Delta t$ which can be implemented using a simple random number generator. We can see such a process for a $10Hz$ sine wave in *Figure 2.7*. The resulting spectrum consists of the single peak of the sine wave at 10Hz and the flat white spectrum of the Poisson process.

The fact that broad frequency peaks and the $\frac{1}{f^\alpha}$ PSD structure, as commonly found in EEG, can not be explained by the irregular firing of an inhomogeneous Poisson process with constant firing rate suggests that it is just a good model for describing the behavior but not the reasons for these distributions. Hence, the peak-width of oscillations commonly seen in EEG can not be explained by a Poisson process but rather variations in the frequencies due to the non-stationarity of the stochastic sources [Jones, 2016].

2.2.1.2 Postsynaptic Potentials

Postsynaptic Potentials (PSP) take place in the dendrites of neurons and are triggered by arriving action potentials on the presynaptic axon. The timescale of IPSPs and EPSPs is different. In particular IPSPs are slower than APs and we can investigate the corresponding spectrum in *Figure 2.5*. As the PSPs are triggered by action potentials as single events, the single synapses share the same firing statistics as the presynaptic

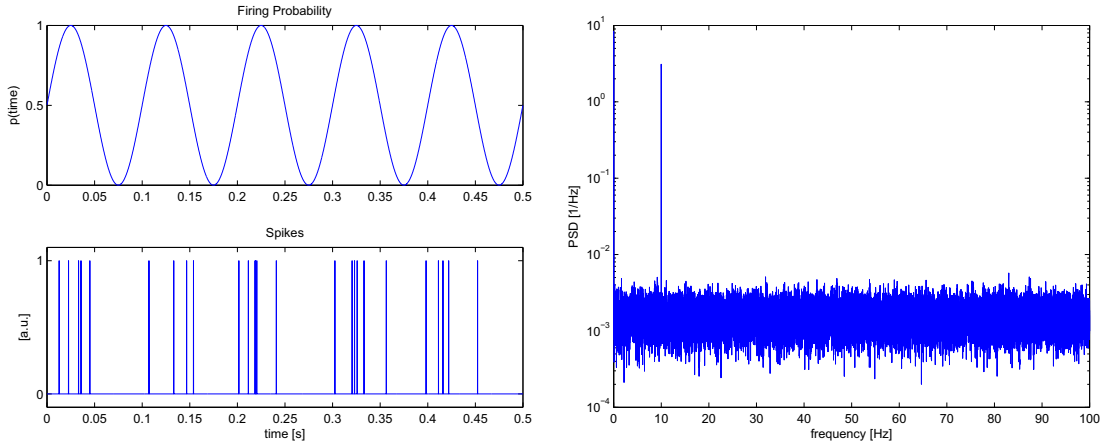

Figure 2.7: A spectrum of an inhomogeneous Poisson spike train driven by a pure sine wave of frequency 10Hz: we can see a flat frequency response with a single peak at 10 Hz superimposed. The single events were modeled as single Dirac delta pulses in this case.

neuron and hence follow the distributions connected to a Poisson process. In order to reproduce a certain structure in spectral behavior, the synaptic input in form of single action potentials has to have this particular structure or lead to it in combination with single cell membrane properties. This implies that a certain rhythmic or $\frac{1}{f^\alpha}$ structure in PSPs is directly related to firing patterns of presynaptic neurons (including local and global feedback). The post-synaptic current (I_{PSC}) is commonly described by two exponentials of the form [Buzsaki, 2006; Gao et al., 2017]:

$$I_{PSC} = C(-e^{-\frac{t}{\tau_{rise}}} + e^{-\frac{t}{\tau_{decay}}})$$

where C is an amplitude constant. A linear mixing of the IPSPs and EPSPs can create various spectral shapes in the PSD depending on the ratio. In [Gao et al., 2017] it was shown that, due to the differing time constants of excitation ($\tau_{rise} = 0.1ms, \tau_{decay} = 2ms$) and inhibition ($\tau_{rise} = 0.5ms, \tau_{decay} = 10ms$), the slope of the power spectrum is influenced by the general level of excitation vs. inhibition (E/I ratio).

2.2.2 Distributed Sources of Spectral Structure

As we have seen in the last section, the single cell firing properties of cortical neurons and their statistics that can be measured by single cell recordings cannot fully explain the LFP statistics measured close to the cells. The spectral structure of single neurons has to be produced by the ensemble statistical structure of the many inputs into each neuron in combination with the active resonant behavior of the neuron and closely connected

interneurons. This is valid for oscillatory peaks but at least partly also for the 1/f statistics, while the connection between input statistics and 1/f noise shape is highly debated.

Additionally, the external currents of near and far fields are further determined by the ensemble activity of distributed neuronal sources. Constructive and destructive interference can amplify and annihilate local source activity in the far field.

We will thus investigate how different phenomena in the spectra may arise by this ensemble activity. We will also introduce a new perspective on it that stems from the non-linear dependence of firing rate and PSPs. This theory can explain the non-sinusoidal shape of oscillations and the related emergence of harmonics in the spectrum through the clipping effects of the firing thresholds.

2.2.2.1 Background Noise and Broadband Structure

Noise is usually characterized by two main statistical qualities: the distribution of amplitudes and the power spectral densities. While the distribution of amplitudes is commonly assumed to be Gaussian, the spectral behavior is a function of the frequency. Noise relevant for EEG measurements can be charecterized by its spectral roll-off and follows the function

$$S(f) = \frac{1}{f^{\alpha}} \tag{2.12}$$

In technical applications, usually three main types of noise are defined based on their spectral structure, while they actually form a continuum based on *Equation* (2.12):

- White noise: $S(f) = \text{const.}$
- Pink noise: $S(f) = \frac{1}{f}$
- Brownian noise: $S(f) = \frac{1}{f^2}$

Pink noise can emerge out of white noise by integration, or - more general - low-pass filtering. The same accounts for Brownian noise. They all are related to basic physical processes: white noise resembles random events like collisions of particles or what we have seen as a random point process in the APs in the last section. Pink noise is often referring to a fractional exponent α between zero and two and is related to flicker noise in electronic devices. Brownian noise charecterized a diffusion or random walk process like Brownian temperature movements of particles.

Integrating the content of the last sections, we can say that actually random point processes can produce any sort of spectral structure depending on the reaction of the system to the single events. In pink noise, the single events are classically thought to be exponential decays and the time constant determines the spectral roll-off [Schottky, 1926].

In EEG, like many biological systems, mostly pink noise with an exponent between one and two, often close to two is found. It is, however, important to mention first that it makes little sense to talk about a $\frac{1}{f^\alpha}$ structure of the EEG in general without discriminating between frequency bands. Many effects described in the literature are limited to certain frequency bands.

There are many different explanations for the measured $\frac{1}{f^\alpha}$ power spectra that are highly under debate [He, 2014]:

Self-Organized Criticality One of the most cited studies in the context of brain, [Linkenkaer-Hansen et al., 2001], investigated long-range temporal correlations in the amplitudes of band-pass filtered EEG and MEG data and discovered different but similar $\frac{1}{f^\alpha}$ dependencies in alpha, mu and beta bands and assessed these to self-organized criticality in the neural network behavior of the brain. Problematic is that many different reasons can actually lead to this sort of behavior, among which self-organized criticality is only one possible explanation, because the PSD is only a second order statistic [He, 2014].

Point Processes As we have shown before, also single point processes like a Poisson process can produce a similar structure depending on the time course of the single event. If the single event has a $\frac{1}{f^2}$ decay like the EPSP, the PSD of the process will have the same structure [Lowen and Teich, 2005], similar to the first ideas of Schottky [Schottky, 1926] about the shot noise described by Johnson [Johnson, 1925]. He assumed a simple exponential relaxation N(t) in the reaction to each single event:

$$N(t) = N_1 e^{-\alpha \Delta t}$$

that leads to a $\frac{1}{f^\alpha}$ spectrum.

E/I-ratio As we have seen in *Chapter 2.2.1.1*, it is the structure of the input and so most probably the network ensemble activity that leads to irregular firing of neurons, not the behavior of single neurons. We know that the firing statistics of single presynaptic neurons is Poisson distributed which leads to the spectrum depicted in *Figure 2.6* for the post-synaptic potentials. For example, [Freeman and Zhai, 2009; Gao et al., 2017] have shown that fitting two exponentials for rise and fall time of EPSPs can explain the spectral response of ECoG data. Also, the background noise as $\frac{1}{f^\alpha}$ seems to be influenced by various tasks conditions, in particular in higher frequencies [Miller et al., 2014]. This could be related to a general increase in presynaptic activation resembling a higher presynaptic firing rate, a local change in excitation inhibition ratios (E/I raio) or even directly originating from clipped membrane potential oscillations as we will see in *Chapter 2.2.2.2*.

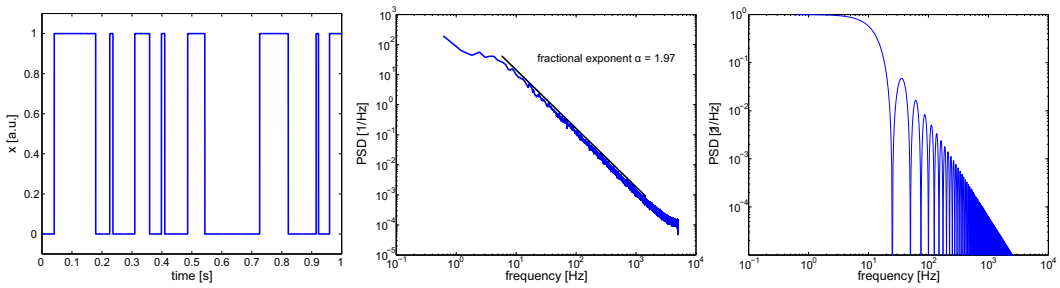

Figure 2.8: A time signal consisting of random rectangular pulses of random size, uniformly distributed over [0 100ms] and its Power Spectral Density (PSD). We can see a clear $\frac{1}{f^2}$ structure without any long range correlations (independent rectangle lengths and positions). C The PSD of a single of the pulses used in A: B is the simple sum of a lot of $sinc^2$ functions with varying frequencies.

Action Potentials Also action potentials influence the structure of PSDs at least locally. Due to the shorter time-pulses and the stochastic firing, the synchrony is quickly lost with distance and action potentials have less influence in the far field. Because of the broad-band peak in the PSD of an AP located around 100Hz, this effect is mostly found as a peak in the Gamma frequency range but the actual influence of firing rate is also found on a broader scale.

Alternative Sources In [Bédard et al., 2004; Bédard and Destexhe, 2008] it was shown that also different aspects of volume conduction and dendritic transmission can lead to similar structures in the PSD of EEG signal in higher frequencies. Opening and closing of ion channels is also expected to produce a low-frequency $\frac{1}{f^\alpha}$ spectrum [Lundström and McQueen, 1974; Musha and Yamamoto, 1997; Novikov et al., 1997] among many others factors. In particular in EEG, even non-neural mechanisms like the sweat cells of the skin produce a temperature and physiological stress dependent alteration of the 1/f structure in the low frequency as well as the electrode impedance influenced by temperature and time [Kappenman and Luck, 2010]. In the very low frequencies, scalp EEG is highly contaminated with single event artefacts from the eyes that lead to a $\frac{1}{f^\alpha}$ spectral structure, which are based on eye movements in form of saccades and changes of line of sight and eyelid closures [Plöchl et al., 2012]. They also have a broadband effect in the Gamma range: micro-saccades have a rate of around 1-2 Hz while the actual pattern is a transient burst in 40-100 Hz related to general muscle artefacts. Muscles activity is promoted by action potentials arriving in muscle fibers, potentially explaining the similarity of the broadband effect of APs on the PSD as in *Figure 2.7* and also explaining one factor, why Gamma activity in EEG is usually difficult to reliably extract.

Excursion: Sinc Function In *Figure 2.8* we can see a small experimental setup: a random sequence of pulses with random length in between 0 and 100ms leads to an almost perfect $\frac{1}{f^2}$ structure in the PSD. Each of these pulses has a Fourier transform of a sinc function of which the maxima follow a $1/f$ spectrum. The simple reason:

$$sinc(f) = \frac{sin(f)}{f} \tag{2.13}$$

For the local maxima $f_{max} = n*\pi$ with $n \in N$ corresponding to the maxima and minima of the sine:

$$sinc(f_{max}) = \frac{sin(f_{max})}{f_{max}} = \frac{1}{f_{max}}. \tag{2.14}$$

This $\frac{1}{f_{max}}$ amplitude spectrum corresponds to a $\frac{1}{f_{max}^2}$ power spectrum. A linear superposition of many of these leads to a loss of the minima and hence a straight $1/f$ structure in the Fourier transform and hence a $\frac{1}{f^2}$ structure in the PSD. The flattening in the PSD at below 10 Hz is due to the limited length of the rectangles (including longer rectangles would shift it to lower frequencies). Random opening and closing of membrane channels could thus reproduce such an event [Lundström and McQueen, 1974; Musha and Yamamoto, 1997; Novikov et al., 1997]. This is an example of an integrated random point process. A completely random event can thus create a $\frac{1}{f^2}$ spectrum which strongly weakens conclusions about self-organized criticality drawn alone from the spectrum.

2.2.2.2 Spectral Peaks

Spectral peaks occur on all scales of spatial resolution from intracellular recordings to scalp EEG. Their origin, however, cannot be explained by the mechanisms of pure feed-forward synaptic transmission. We will look at different mechanisms that can produce peaks in the PSD including distributed but - for the occurrence of harmonics - also local mechanisms.

For an oscillatory behavior, a dependence on the recent past in shape of some feedback of activity has to be present [Buzsaki, 2006]. Actually, to provide stable oscillations in a stochastic system with non-stationary noise, a certain non-linearity is needed in order for the oscillation not to diverge. These non-linearities are found in form of the firing thresholds and the maximum firing rates of neurons, often described by a sigmoid function.

Concerning the different peaks commonly found in neurophysiological recordings, there is a large debate about their role and distinctions. Classically, frequency bands are mostly separated into alpha (8-13Hz), beta (15-30Hz), gamma (>30 Hz) and theta band (4-6 Hz) in the awake conscious state. Additionally delta (1-4Hz), sleep spindles (10-20Hz) and slow rhythms (0.05-1Hz) occur in sleep and similar states of (un-)consciousness. The focus throughout this thesis will be on the awake conscious human.

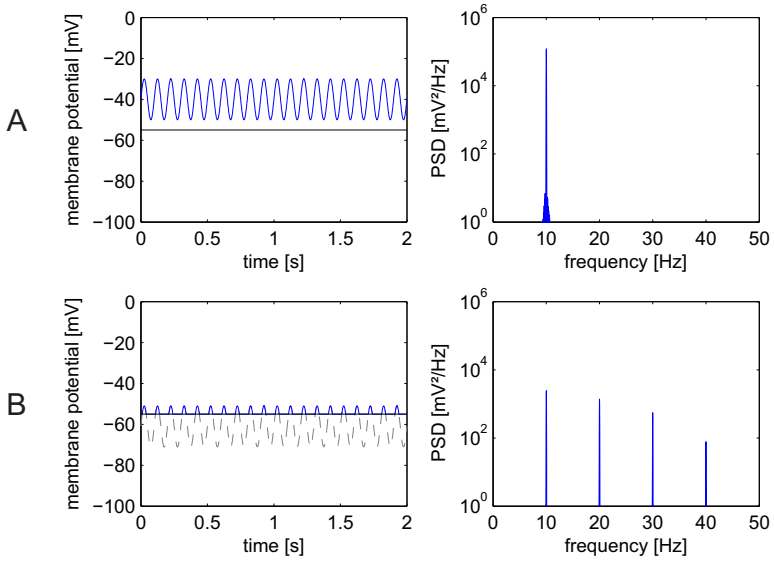

Figure 2.9: Harmonic (**A**) and relaxation oscillators (**B**) as action potential outputs in a neuron produced by super-threshold and around threshold oscillations of the membrane potential: a simple toy model to explain the emergence of harmonics. By clipping the alpha oscillation in **B**, the PSD at the fundamental frequency (10Hz) is decreased, while harmonics appear. **A** can be basically interpreted as a version of **B** with a constant excitation term added.

The separation of bands has its root in the occurrence of different rhythms in different states over different areas [Wróbel, 2000] but is in general not clearly visible and straight-forward detectable in real-world data. In particular, the alpha rhythm introduces harmonics due to its non-sinusoidal shape, which mix with beta and gamma peaks, see *Section 2.2.2.2*. Also, the fact that local stochastic resonances can in general synchronize to the harmonics of an external driver, makes the phenomena indistinguishable as they influence each other.

Alpha and theta are thought to stem mainly from sub-cortical structures: while alpha is related to the cortico-thalamic feedback network, theta is related to hippocampal and hippocampal-cortical activity. Beta and gamma oscillations are commonly linked to more local connections within different cortical and gamma also sub-cortical areas. A generic framework based on neural mass models can explain this by simple conduction delays depending on distance and connectivity.

Oscillations and Stochastic Resonance In [Buzsaki, 2006] the concept of harmonic and non-harmonic oscillators in order to investigate the oscillations in the brain was introduced. Most oscillators in the brain are non-harmonic of a pulsatile type describing

their rhythmic discharge in shape of spike trains of neurons.

The two phenomena (harmonic and non-harmonic) are not separable but form a continuum depending actually on the average membarne potential and the time delay of the feedback. In the brain, we find mainly non-linear dependence of the firing rate on the membrane potential with firing thresholds at minimum and maximum excitation levels. We can transform a harmonic (sinusoidal oscillator) into a relaxation (pulsatile) oscillator with long silence phases by simply clipping the output which is likely to happen with around-threshold oscillations in the brain. Most biological oscillators are relaxation types including heartbeats, respiration, walking and hormone secretion [Buzsaki, 2006].

In neural networks like the brain, a myriad of feedback connections exist between different neurons, on large-scale, sub-networks and close to single neuron perspectives. These include inhibitory interneurons establishing self-feedback to a single neuron and its neighbors but little long-range connections with other cortical areas [Deco et al., 2008]. While most long-range feed-forward connections are found to be dependent on chemical synaptic transmission, in particular fast local inhibition in the gamma range occur to be based on electrical synapses based on gap junctions between interneurons [Lefler et al., 2014].

Membrane Potential Oscillations Subthreshold membrane potential oscillations (SMPO) have been found in slices of different areas of the cortex, where neuronal membranes show resonances at certain frequencies [Lampl and Yarom, 1993; Puil et al., 1994; Pike et al., 2000; Pedroarena and Llinás, 1997]. Through experimental blocking of different ion channels, the conclusion emerges that the resonance is based on their interplay. In brain slices in-vitro, the ion channels of Na^+ and Ca^{2+} have been found to influence resonances in lower but also in higher gamma frequencies [Puil et al., 1994; Pike et al., 2000; Pedroarena and Llinás, 1997]. Most of the studies on SMPOs were done in slices of neurons with remaining network connectivity leaving an open question whether the blocking of the channel actually inhibits a neuronal resonant mechanism or only inhibits the resonance of the network. [Lefler et al., 2014] have shown that direct electrical coupling of neurons in form of gap junctions can lead to these SMPOs and hence supports the network generation hypotheses while it is out of question that properties of ion channels can at least enhance or attenuate reactions at certain frequencies, depending on their time constants [Dickson et al., 2000].

Traveling Waves and the Dispersion Relation [Nunez and Srinivasan, 2006] investigated a theoretical basis for standing and traveling brain waves in human EEG that is based on a dispersion relation to explain spatio-temporal dynamics. The further developments of this approach are handy at explaining the resting-state activity of the brain, the occurrence of certain frequencies and their spatial spread as modes of the basic dispersion relation. [Deco et al., 2008, 2011] were some of the first to prove that slow

fluctuations (<0.11Hz) in the blood-dependent deoxygenation level (BOLD) of fMRI [De Luca et al., 2006; Damoiseaux et al., 2006] and linked to EEG [Mantini et al., 2007] can be modeled by stochastic local oscillators and their delayed coupling with other cortical areas. The assumed conductance velocities were around $1.5m/s$. They particularly could show, how a sparsely connected network with a lack of long-range inhibitory connections can explain in-vivo measured correlations between activities in different areas. This can also be related to local 40Hz (Gamma) oscillations.

Recent studies on resting state functional data of the brain have revealed resonance phenomena on different scales that are topologically sorted and can be related to modes of the Laplace operator on the functional connectivity of the brain [Robinson et al., 2016; Atasoy et al., 2016, 2018]. Based on diffusion tensor images (DTI), a technique related to magnetic resonance imaging, so-called Connectomes estimate functional connectivity between different brain regions. These can then be used to build macroscopic field models of the brain that are shown to resonate in certain modes [Robinson et al., 2016]. These modes are based on the spatio-temporal coupling of the activity $\Psi(x,t)$ of different brain areas x in time t. This can be described by Laplace's equation:

$$\Delta\Psi(x,t) = F(x,t) \tag{2.15}$$

linking time and space by the average neuronal coupling in combination with conduction delays and the external input in the function $F(x,t)$. We can finally transform the dispersion relation into a Helmholtz equation where the eigenvalues of Laplace's operator Δ are the principal modes [Atasoy et al., 2018]:

$$\Delta\Psi_k = \lambda_k\Psi_k \tag{2.16}$$

Similar models can also explain local emergence of those rhythms in cortical areas due to the feedback between pyramidal, stellate cells and inhibitory interneurons in combination with different synaptic time constants of excitation and inhibition[David and Friston, 2003].

State-of-the-art research also proposes a strong roll of Olygodendrocites, a type of Glia cells, in the manipulation of transmission (plasticity) by a change in myelinization of the axons of neurons. The myelinization is important for the timing and the amplitude of the communication between cells [Pajevic et al., 2014; Fields, 2015]. Thus a fine tuning of the large-scale oscillations could be promoted by such a mechanism an enhance or decrease the propagation of certain frequencies.

Exemplary Rhythms: We will now look at two important examples of rhythmic activity stemming from sub-cortical structures that we also commonly find in EEG measurements. They can be related to the described resting state oscillations by the structure of their functional connectivity:

Thalamo-Cortical Rhythms: In the Thalamus of awake and conscious humans two main rhythmic frequencies are found: alpha/mu/tau rhythms (8-12Hz) and ultra-high frequency oscillations (300-600Hz), while there are also different sleep related waves and spindles not discussed here [Buzsaki, 2006]. The source of the alpha-like rhythms is highly under debate and the emergence is commonly related to local pacemakers as well as thalamo-cortico-thalamic feedback loops [Suffczynski et al., 2001]. The ultra-high frequency is most likely related to local inhibitory feedback networks, which will be further discussed below. Functionally, we find oscillations in the 10-Hz range to be relevant for sensory-motor activity (mu) [Pfurtscheller and Neuper, 2001], visual (tau) [Sauseng et al., 2005] and auditory attention (alpha) [Lehtelä et al., 1997] and hence also within the corresponding cortices.

Hippocampal Rhythms: The Hippocampus of awake humans has two main rhythms: theta and sharp wave gamma rhythms [Colgin, 2016] but also other rhythms like a 3-Hz oscillation have been found [Matsumoto et al., 2016]. The origin of the sharp-wave ripples in rats has been accounted to local inhibitory networks similar to those of the thalamus and data supports that similar mechanisms could be present in humans [Le Van Quyen et al., 2008]. The theta rhythm of Hippocampus is related to memory [Kirchner, 1958] and the same rhythm in fronto-central locations of EEG is usually related to working memory activity [Gevins et al., 1998].

Gamma and high-frequency spike bursts The Gamma and higher frequency firing and bursts are directly related to the firing frequency of single neurons also seen in classical Hodgkin-Huxley membrane models. The so-called refractory state of neurons, related to the hyper-polarization caused by potassium channels, is a time period of $3ms$ on average with blocked excitability of action potentials. The maximal firing frequency during classical excitation based firing is thus around $f_{max} = \frac{1}{3ms} = 333Hz$. But, depending on the location of the specific neuron within the brain and cell layer, the neurons are tuned to different spectral behavior for the same input. Fast spiking inhibitory interneurons can reach instantaneous firing frequencies up to over 800 Hz while making up about one third of neocortical interneuronal population [Wang et al., 2016]. Under normal conditions the preferred firing frequency of pyramidal cells is rather in the range of 50-80 Hz, while for these cells maximal rates of around 200-300Hz are found [Cardin et al., 2009].

We have seen that the firing rate of neurons modeled as Poisson point processes do not produce peaks at the relevant frequencies (see *Figure 2.6*) but only a reproduction of the shape of the Action potential's PSD with added white noise. While in this figure, the AP of a pyramidal cell was modeled, fast-spiking interneuronal APs consist of much more narrow peaks [Wang et al., 2016] and hence the spectral characteristics differ. The shorter time constants will lead to a shift of the main activity to higher frequencies. Nevertheless,

the direct spectral influence of APs is rather broadband in both cases. Hence, temporal ripples (or equivalently spectral peaks) need some sort of (local) feedback in order to consist. Peaks in the lower gamma range are usually found to be spectrally broad and spatial focused. They are mostly thought to be related to an asynchronous increase in firing of larger assemblies of neurons [Miller et al., 2007], most probably pyramidal cells.

In the gamma band of rodents (here 20-80 Hz), the main source of spectral peaks of cortical origin is related to local feedback by inhibitory interneurons [Cardin et al., 2009], while we will see later that in particular in the range below 50Hz, also harmonics of alpha and other low frequency rhythms play a role. The gamma band oscillations need to include synchronous firing to produce a far field and are hence related to so-called fast-spiking interneurons connected by gap junctions to local excitatory neurons synchronizing them [Hasenstaub et al., 2005; Wang and Buzsáki, 1996]. They are functionally related to an increase in evoked response and focused attention [Fell et al., 2003], meaning the synchronicity in the network spiking activity seems to be related to afferent excitation and, hence, the network input to these areas [Bair and Koch, 1996; Nadasdy, 2010].

In particular in thalamic neurons, rhythmic ultra-high frequency spike burst (300-700Hz) have been found [Buzsaki, 2006; McCormick and Pape, 1990]. An extension of the Hodgkin-Huxley model is necessary to account for these spike bursts due to its relation to afterhyperpolarization caused by slow Calcium currents. Oscillations in this frequency range are also found in and over somatosensory-areas [Curio et al., 1994; Baker et al., 2003; Fedele, 2014]. Interestingly, this kind of spike bursts are limited to a maximum of 10Hz modulation rate in single neurons, as they need an inactivation period of longer than 100ms to occur [Sherman, 2001], which is again a hint towards the necessary ensemble activity to create higher-frequency patterns.

The synchrony of the spikes in different neurons are promoted by local networks of fast-spiking inhibitory interneurons, also for ultra-high-frequency spike-bursts [Draguhn et al., 1998]. The highest firing frequency of feed-forward neurons like pyramidal cells is found in the range of 200-400Hz meaning that oscillations measured at higher rates could be traveling waves of spike trains promoted by local inhibitory inter-neuronal networks [Draguhn et al., 1998] or directly the firing rates of those fast-spiking cells. Cortical high-frequency as well as ultra-high-frequency spike bursts are most likely related to highly synchronous action potentials in the axons of thalamic afferents to the cortex and the reaction to them in local feedback [Ozaki and Hashimoto, 2011].

[Nadasdy, 2010] suggested a model that predicts synchronous firing of many neurons when sub-threshold oscillations are modified to just cross the threshold which could also be related to afterhyperpolarization effects in the shape of a relaxation oscillator. This model includes the phase-locking of local high-frequency oscillations to low-frequency oscillations like alpha and theta. The synchrony between the spikes of different cells at the higher frequency often measured [Curio et al., 1994; Baker et al., 2003] can be explained by an additional inhibitory interneural network. An interesting point is that

high and low frequency oscillations do not need to be synchronous to explain these effects.

Higher frequency oscillation are thus most probably directly related to firing rates and their local inhibitory feedback, the synchronization of which quickly decreases with distance. The very coarse neural mass models that can explain basic oscillations by dispersion relations inherently incorporate high-frequency oscillations by modeling inhibitory feedback only as fast and local.

Non-Sinusoidal nature of signals Neural signals are known to be non-sinusoidal and changes in the oscillatory shape seem to play an important part in diseases like Parkinson's [van Dijk et al., 2010; Cole and Voytek, 2019; Schaworonkow and Nikulin, 2019]. Let us get back to *Figure 2.9*: we can see how clipping of a sine wave introduces harmonics in the frequency spectrum and hence in the PSD. We find different aspects of clipping: the symmetry and the softness of clipping. While the symmetry influences the occurrence of odd and even harmonics, the softness of clipping determines the spectral roll-off. We find clipping and soft-clipping in the sigmoid dependence of the firing rates on the synaptic activity. Hence, interpreting the individual spectral peaks in the PSD independently makes in general no sense. We will now look at a model of a single neuron mechanism that produces non-sinusoidal harmonics even when driven with a pure sine wave.

In *Figure 2.10* we can see the firing rate of neurons depending on the current injection relative to the firing threshold current into the dendrites. The different transfer functions of neurons can be modeled with the Connor-Stevens model [Connor and Stevens, 1971], which is an extension to Hodgkin-Huxley membrane models. We can see a sudden jump in firing rate to a maximum firing rate which, here describes the afterhyperpolarization spike bursts described in the last sections, while in usual tonic firing, the increase is more constantly. In all of these models, we can see a non-linear compression of input, which in fact does lead to described effects of (soft-) clipping and hence introduces harmonics into the spectra. This can in general be described by a sigmoid function.

Hence, we find a sigmoid shape of the frequency depending on the input as the main root of non-linearity in neural transmission. It can be included into the general dispersion based wave model of neural activity [Nunez and Srinivasan, 2006; Deco et al., 2008; Robinson et al., 2016; Atasoy et al., 2016, 2018] in order to link micro- and macroscopic perspective of brain function [Jirsa and Haken, 1997].

In *Figure 2.9* we have seen how simple clipping of perfectly sinusoidal signal leads to the emergence of harmonics in the amplitude spectra. *Figure 2.11* shows, how the non-linearity in the frequency dependence of firing on the membrane potential leads to a soft-clipping or compression around the minimal and maximal firing thresholds. These non-sinusoidal in the time domain lead mainly to the occurrence of spectral peaks at the harmonics (multiples) of the base frequency. If we simulate a more realistic alpha

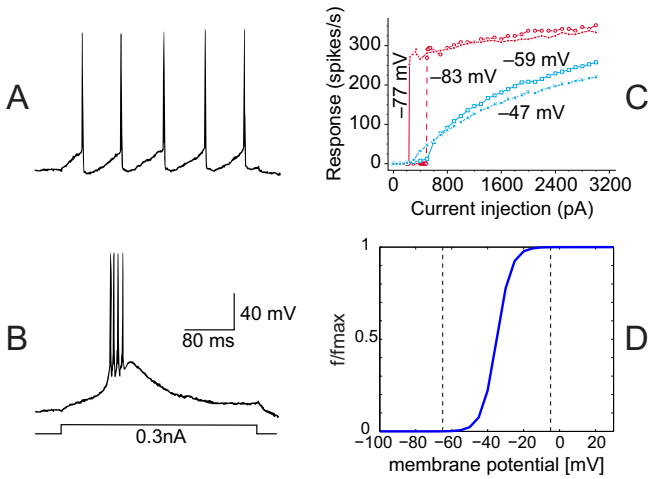

Figure 2.10: Firing rates of neurons: **A** Tonic firing **B** afterhyperpolarization spike burst **C** Firing rates of both processes and**D** logistic model of firing rate commonly used in artificial neural networks. In **D**, the dashed lines depict the practical soft-clipping threshold for an input-output relationship. Note that the recordings in **C** were done under patch clamp recordings fixing the membrane voltage The blue lines correspond to tonic while the red to afterhyperpolarization spike bursts. Reprinted with permission from Copyright 2001 Elsevier. **A**,**B**,**C** reprinted from [Sherman, 2001], Copyright (2001), with permission from Elsevier and [Zhan et al., 1999], copyright 1999 The American Physiological Society.

peak, apply the sigmoid compression and calculate the power spectra, we find that the occurrence of other peaks, in particular that of even harmonics like beta, can depend on these effects. Hence, even a seemingly independent occurrence or modulation of single peaks can be reduced to the effect of a single variable, meaning they are not necessarily independent as often assumed [Wolpaw and McFarland, 2004; Nikulin and Brismar, 2006; Hari and Salmelin, 1997]. Also, [Wolpaw and McFarland, 2004] actually made their conclusions on the ability to differential control of left hand area alpha vs. right hand beta, which mixes up the effects of frequency and spatial location. The data also shows a co-modulation of alpha and beta and hence strongly weakens the conclusions drawn. In [Nikulin and Brismar, 2006; Hari and Salmelin, 1997] the difference in scalp patterns is taken as a hint towards distinct mechanisms, while mostly an overlap of the larger spatial scale alpha with beta is present. This does not proof distinct processes but could still mean only partly modulation of harmonics, while further (local) resonances additionally influence the spectral powers and phases.

In *Figure 2.12*, we see an inhomogeneous Poisson process driven by a single alpha peak

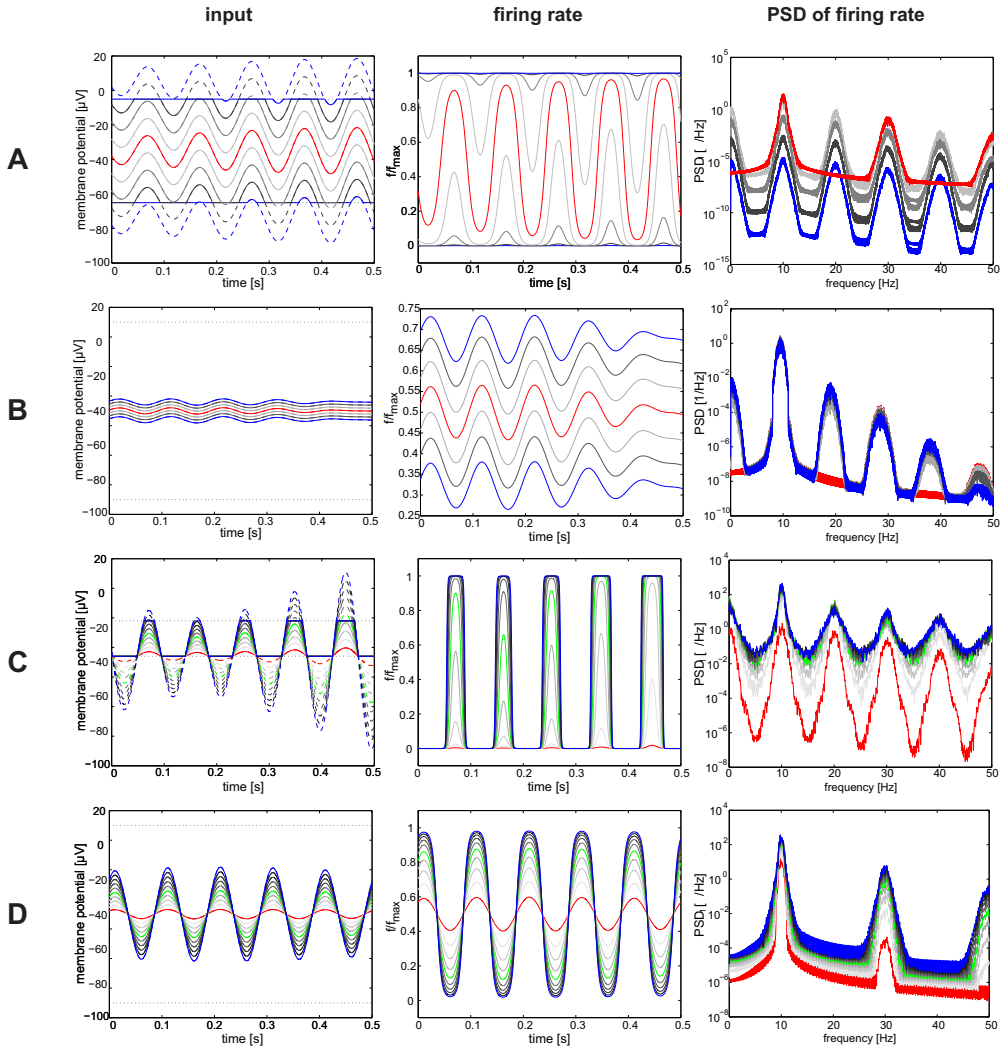

Figure 2.11: Sigmoid compression of an alpha oscillation by various parameters: **A** the asymmetry/offset in large, **B** the asymmetry/offset in detail, **C** the amplitude of 'subthreshold'/asymmetric alpha, **D** the amplitude of symmetric alpha. The decay of the harmonics' peak amplitudes and their dominance depends on the amplitude of the base oscillation. The individual amplitude of each peak depends on the asymmetry/offset. We can also observe that the peakwidths increase with frequency. No independent noise was added, hence the background noise comes from the simulation process of the alpha peak and is thus modulating with amplitude. The spectral roll-off of noise in the input is thus also modulated with its amplitude by the emergence of harmonics.

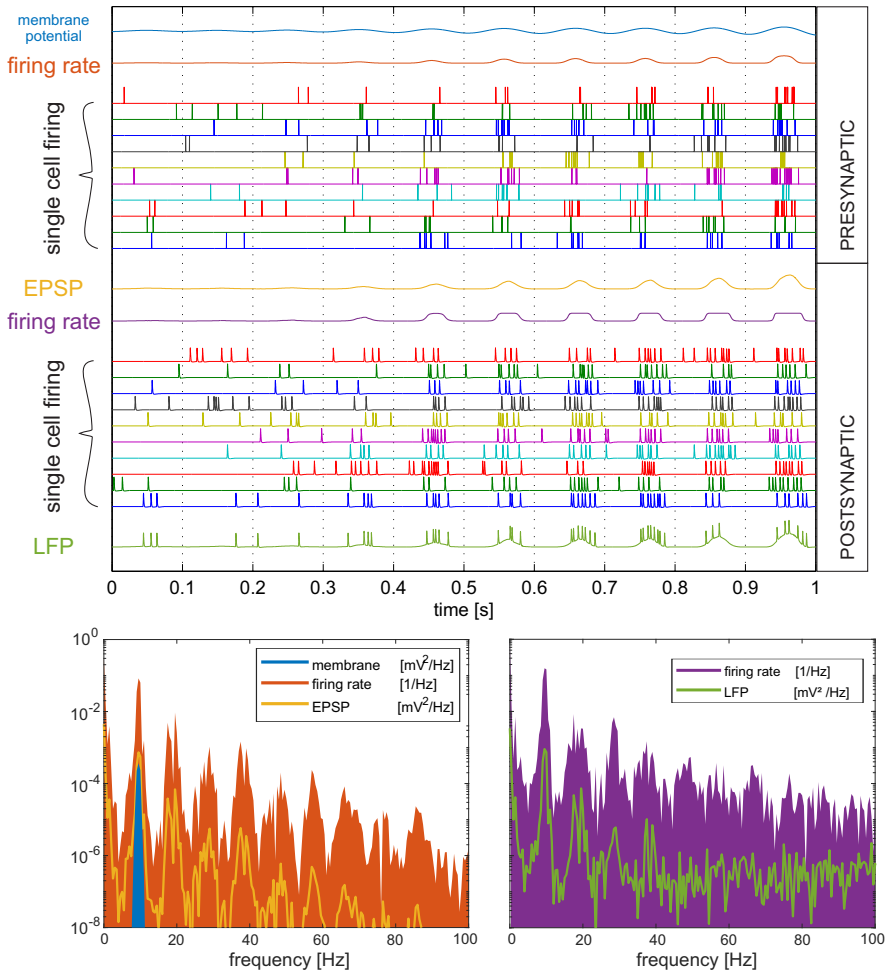

Figure 2.12: A simple model of synaptic transmission: an inhomogeneous Poisson process with an alpha subthreshold membrane potential oscillation leads to single event APs arriving at the synapse. These are transformed via a sigmoid function into a time-depending normalized firing rate (f/f_{max}) for a second inhomogeneous Poisson process in the post-synaptic neuron. The local field potential (LFP) shows both, the alpha rhythms and the gamma bursts on top. This is also a model for cross-frequency interactions like theta-gamma phase-amplitude coupling. You can see the small delay caused by the EPSP at the synapse. Not included here are the stochastic resonances (e.g. at gamma frequencies) of the postsynaptic cells that lead to further peaks in the spectra. The scaling is arbitrary.

that can explain some of the measurements usually obtained. The design is strongly simplified and transmission delays are neglected, but this resembles in principle the output of the thalamus onto a neocortical layer. The oscillation is first transformed into a firing rate using a sigmoid function and then into a time signal using an inhomogeneous Poisson point process in combination with single APs as events. At the neocortical layer, the APs are transformed into EPSPs in a second set of neurons which trigger a new set of action potentials based on the same mechanisms as in the Thalamus. The resulting signal consists not only of the base frequency, but also harmonics and broadband power increase related to the APs. Inhibition is not taken of care in this simple example.

Strong coupling of in particular alpha and beta amplitude and phase are generally found, suggesting a common mechanism of creation [Carlqvist et al., 2005; Nikulin and Brismar, 2006]. This could be at least partly explained by the effects of non-linear neuronal firing rates.

Another factor is the width of the spectral peaks, which could be also an indication of the harmonicity. The width of the peaks in alpha is usually about half as wide as the one in beta, while taking the double frequency values of alpha explains this. A varying base frequency leads to double the variance in the first harmonics, triple in the third and so on.

In *Chapter 2.2.1.1* we have seen that the firing of single neurons is unlikely to cause the imprecision in the oscillation frequency and so the width of the peaks in the spectra. Hence, three possible sources are reasonable: either the thalamic source of the oscillations consists of a horde of single oscillations with different frequencies, they are relayed by stochastic resonators with varying frequencies or the whole oscillation is based on long-range feedback loops of thalamo-cortico-thalamo structure with varying transmission delay. The latter goes well in line with the neural field theories of average connectivity and transmission speed leading to a variety of resonances that need to be diffuse on their spatial boundaries of the basic modes in order to coexist.

[Klimesch, 2012] and others strongly suggested alpha oscillations to be a distinct mechanism of active inhibition/suppression to enhance focus to non-inhibited areas, while this could also be explained by an area of the cortex moving to a more local and thus higher frequency mode in neural field and similar theories or the creation of harmonics. [Haegens et al., 2011] found a increase in the firing rate in LFPs in monkeys that was coherent with a decrease in alpha and an increase in beta power. While they concluded that alpha is thus responsible for active inhibition/suppression, these findings go also inline with hypotheses of the creation of harmonics due to a general increase in an excitation level from a balanced mean firing level of excitation. This could be additionally selectively amplified by stochastic resonances of the receiving cells and their network.

Summarizing the most plausible effects based on the theory of this chapter, many factors contribute to the shape of the PSD. While the effects related to E/I-ratio and population firing rate seem more local and thus are rather measured intracranially, EEG seems mainly influence by mass PSPs. LFP and ECoG spectra include all effects: de-

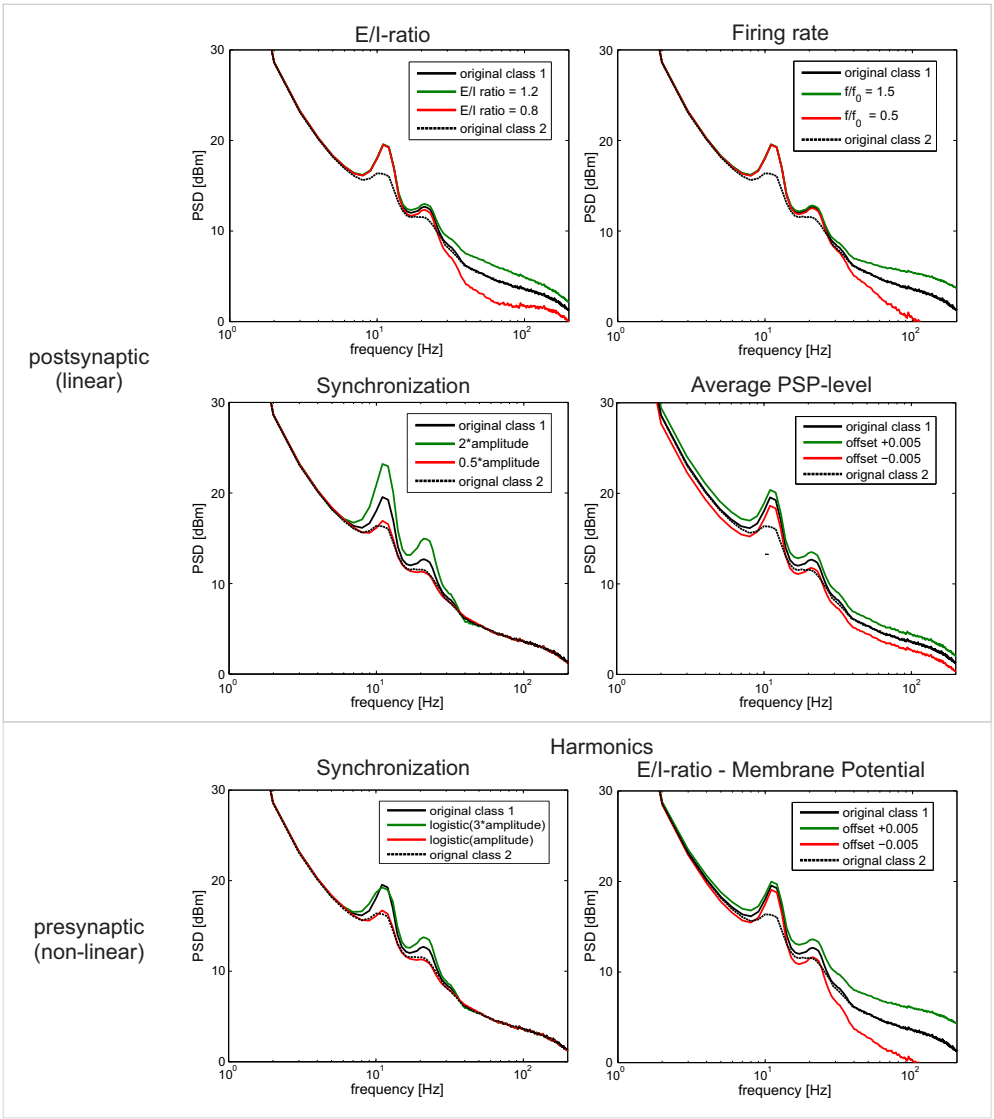

Figure 2.13: Different hypothetical changes in the spectra due to different mechanisms in comparison to real EEG data: a change in E/I-ratio changes the spectral roll-off in frequencies above 15 Hz, a change in firing rate the spectral roll-off in even higher frequencies. A change in synchronization of oscillations increases or decreases peaks and their harmonics, while the same presynaptically changes in the non-linearity changes the spectral roll-off in the harmonics and also the amplitude of the base frequency. The average PSP level leads to an overall increase/decrease of the spectrum while presynaptically, it has an effect on the spectral roll-off for $\frac{1}{f^{\alpha}}$-shaped spectra.

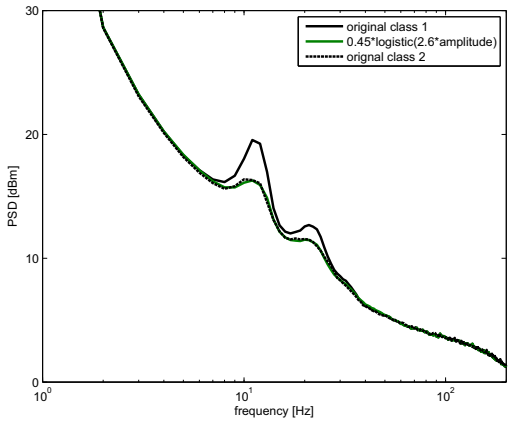

Figure 2.14: Investigating the most plausible explanation for the class difference in EEG between left and right hand motor imagery based on the assumption of beta being a harmonic of alpha: a linear increase in the presynaptic synchronization leading to a nonlinear amplitude change in the firing frequency of presynaptic action potentials of which the mean field amplitude is decreased by desynchronization across neurons.

/synchronization, E/I-ratio and firing rate. E/I-ratio and firing rate are challenging to distinguish as the spectral change is similar and also functionally interrelated, because a higher excitation level leads to higher firing rate. However, the directly increasing influence of APs with firing rate is happening rather in higher frequencies due to the short time constants of the process. The postsynaptic average PSP-level in the mean field of the cortices has not neccessarily a large role, because a general shift of the spectrum is not observed. Harmonics, on the other hand, are capable of explaining changes in spectral peaks that are not independent from each other. A mixture of all effects is in general most likely and challenges interpretations. It depends on the spatial resolution of the measurement technique, which effects can be observed. The logistic firing probability is an approximation and not neccessarily very realistic. Also a change in the firing properties of single neurons can have similar effects.

The class differences are most likely explained by a presynaptic synchronization with a postsynaptic desynchronization, while other combinations like independent alpha and beta oscillations with individual synchronization/desynchronization or a change in the firing rate dependence on the membrane potential may lead to similar spectra.

Cross-freqeuncy Interactions A very deep result is that a perfectly sinusoidal oscillation at alpha frequencies around the firing threshold of neurons in combination with a non-linear relationship between the membrane potential and the firing rate can lead to a phenomenon we often observe: high-frequency spike bursts on top of lower-frequency

oscillations. This is observed often in intracranial recordings, but it can also be extracted from EEG with special equipment and experimental paradigms [Curio et al., 1994; Baker et al., 2003]. The offset of the oscillation from the threshold modulates the output from no output via oscillations to tonic firing of a single neuron (see *Figure 2.11*).

Coupling of the different oscillations was recently further investigated [Jensen and Colgin, 2007], it describes the interaction of different frequencies' phase and power which can be explained by non-linear clipping of the synaptic input by the cells introduced within this thesis. Actually the multiplication in the time domain is related to a convolution in the frequency domain. Hence, a basic sine wave is transformed by a sinusoid and its output is multiplied with a higher frequency modulation. This basically resembles a phase power relationship but also a power-power relationship due to the firing rate being influenced by the sine wave amplitude.

2.3 Chapter Summary

The spatial separation of charges due to the synaptic currents and similar events leads to a current flow along the dendrites inside and outside the neuron due to the limited passive conductivity of the membrane. The same accounts for APs but spatially on a smaller scale. The equilibrium of charges is reached by return currents along the dendrites and axons inside and outside of the cells. The outside current is termed volume conduction and leads to a potential distribution across the whole head depending on the neural activity. The neural sources can mostly be approximated by a dipolar field and resulting field equations are solved in electrostatics leading to a linear voltage-current relationship. The external field increases with spatial scale and the synchrony of the current sources. This leads to the conclusions that EEG consists predominantly of postsynaptic potentials of neocortical pyramidal cells.

Various factors contribute to the spectral shape of the electrical signals produced during neural activity. These factors range from frequency preference of single neurons via network based synchronization and oscillations up to being simple harmonics of their base frequencies due to their non-sinusoidal nature. An interplay of all these factors suggests that oscillators on all levels can lock in to necessarily occurring harmonics and resonate in spectral and spatial modes of the base frequency, while a second local inhibitory network synchronizes higher frequency spiking in the gamma range up to around 200Hz. The inhibitory networks themselves incorporate fast-spiking interneurons that can have higher firing frequencies up to around 800Hz and also produce high frequency temporal patterns most probably as a result of wave propagation within their network.

Constructive and destructive interference in the linear superposition of source currents due to synchrony and asynchrony lead to a predominance of different spectral peaks at different levels of spatial resolution, while the large spatial scale EEG signals mainly deliver the low-frequency regimen.

Chapter 3

Methods for Electrophysiological Signal Analysis

3.1 Imaging and Recording Methods

Neuroelectrical signals can be measured in different levels of spatial resolution and either directly in the electrical domain or magnetic fields resulting from volume current flows. Non-invasive recording techniques like Electroencephalography (EEG) [Berger, 1931] or Magnetoencephalography (MEG) [Cohen, 1968] exhibit comparatively low signal-to-noise ratio (SNR). In contrast, invasive techniques like Electrocorticography (ECoG) or the measurement of local field potentials (LFPs) [Henrie and Shapley, 2005] have higher SNR but are not (yet?) applicable to standard experiments on healthy subjects or even everyday applications. Measuring invasively increases spatial resolution due to decreased spatial smearing and lower interference from artefacts of eyes and muscles.

Imaging methods like Magnetic Resonance Imaging (MRI) [Pykett, 1982], Electrical Impedance Tomography (EIT) [Barber and Brown, 1984; Holder, 2005] are mainly anatomical imaging methods, while near-infrared spectroscopy (NIRS) is intended to estimate the light spectral properties of materials. But, the blood oxygenation as a metabolic effect of neural activity, can also be detected in functional MRI (fMRI) [Belliveau et al., 1991] and functional NIRS (fNIRS) [Jobsis, 1977]. Hence, they serve as an indirect measure of neural activity through the Blood-Oxygenation-Level-Dependency (BOLD) on neural activity. Local conductivity changes due to the change of ion concentration leads to EIT being also sensitive to neural activity [Aristovich et al., 2016]. We will however focus on their use as anatomical imaging methods.

3.1.1 Invasive Methods

3.1.1.1 Intracellular Recordings

Almost exclusively performed in animals models, intracellular recordings are the unique method to directly measure the sources of extracellular fields: fluctuations in the membrane potentials [Buzsáki et al., 2012]. This includes action potentials and post-synaptic potentials among more specific approaches like single ion channel patch clamps.

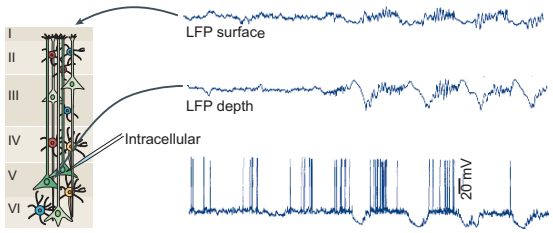

Figure 3.1: An example of simultaneous intra- and extracellular recording: The intracellular recordings clearly reveal action potentials and hyper-polarizations in the membrane potential, while the AP spikes are visible extracellularly but with strongly decreased amplitude and a loss of temporal focus. Adopted with permission from [Buzsáki et al., 2012] and [Contreras and Steriade, 1995], copyright 1995 Society for Neuroscience

3.1.1.2 Local Field Potentials

Local field potentials (LFPs) are measured in the proximity of neurons within the cortex (see figure *Figure 3.1*) [Henrie and Shapley, 2005]. Usually needle-like pins with several electrodes - termed tetrodes for four - are inserted into the cortex and placed close to cells in the targeted cortex area with the help of audio-guiding. The electrodes are positioned such that the geometric source location can be well estimated and more importantly the signals of different cells can be distinguished. Local field potential pick up the electrical field of all surrounding cells and PCA and ICA are usually applied to separate the different sources. The measurements mainly catch EPSPs and APs of the closest cells, as the amplitude of the in first approximation dipolar field decreases quadratically with radius (compare the equation in *Section 2.1.1.2*) [Buzsáki et al., 2012]. In *Section 2.2.2.2* we have seen how cross-frequency interactions like alpha or theta coupling with the gamma range occur. These are often measured in LFPs.

3.1.1.3 Electrocorticography

Electrocorticography (ECoG) is measured directly on the cortex surface inside the head leading to a spatial resolution between LFPs and EEG. As discussed in detail in *Chapter 2*, ECoG provides a larger coverage across brain tissue while its sensitivity to local signals is much higher than in EEG, as EEG suffers from the effects of skull and CSF as a intermediate isolation and shielding layers between sources and sensors [Miller et al., 2007]. Hence, we can observe mean field network activity on a smaller spatial scale than EEG while the analysis is mostly limited to epileptic or similar patients. This is maybe the biggest drawback, as most research done in this field is based on data of patients with neurological disorders under the influence of medication during surgeries, which alters brain function and the characteristics of the signal such as its power spectral density

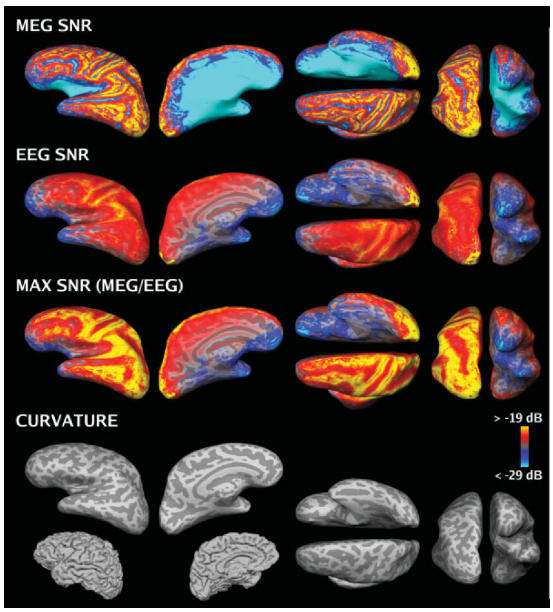

Figure 3.2: SNR comparison between MEG and EEG: the maximum SNR of EEG is in general lower than that of MEG while it is higher away from the surface. MEG has a higher SNR close to the surface, which decreases strongly with centricity. Adopted with permission from [Goldenholz et al., 2009], copyright 2008 Wiley-Liss, Inc.

(PSD). However, ECoG drastically improves spatial resolution in the frequencies from beta upwards and is hence an irreplaceable tool for investigating brain activity.

3.1.2 Non-Invasive Methods

3.1.2.1 Electroencephalography

Electroencephalography (EEG) [Berger, 1931; Nunez and Srinivasan, 2005] is the most widely used recording technique for neural activity in humans because of its low cost, non-invasive nature and ease-of-use. Electrodes are placed on the surface of the head, and the potential differences between different locations can be measured. This gives a spatio-temporal image of the electrical activity with a high temporal but a low spatial resolution. The spatial resolution can be increased by spatial filters that counteract the strong spatial smearing effects of in particular CSF and skull onto the volume conduction. Biophysical models of the volume conduction can additionally help to localize the sources.

EEG, however, also (or mainly) catches the eye movements and lid closures, Electroocolugram (EOG), the activity of muscles, Electromyogram (EMG), including the

huge heart muscles, termed Electrocardiogram (ECG) among pulse wave related effects (Mayer wave) and other biological or electrical noise caused by the measurement equipment (electrodes/amplifiers). The main challenge is that many of the mentioned sources have a higher amplitude in the EEG measurements than the neural activity.

3.1.2.2 Magnetoencephalography

Magnetoencephalography (MEG) measures the magnetic field resulting from the neuronal volume currents of neural activity within the brain [Cohen, 1968; Häamäläinen et al., 1993; Hari and Salmelin, 1997]. The measurements are thus not effected by the shielding of CSF and skull while the low SNR restricts measurements to superficial sources. Additionally, technical details additionally result in a sensitivity to tangential dipoles, while radial dipoles are hardly measurable [Eulitz et al., 1997].

3.1.2.3 Magnetic Resonance Imaging

Magnetic Resonance Imaging (MRI)[Pykett, 1982] is a technique that is based on the magnetic spin polarization of predominantly water molecules in the body and can hence be used to receive anatomical images with high contrast of soft tissue as shown in figure *Figure 3.3*. Additionally, as the magnetic properties of blood depend on the oxygenation level, opens the door to the detection of metabolic effects of neural or other activity in functional Magnetic Resonance Imaging (fMRI) [Ogawa et al., 1990]. Further, Diffusion Tensor Imaging (DTI) - a special recording technique in MRI - can estimate the location, orientation and density of white matter fiber tracts which have an effect on the anisotropy of volume conduction [Basser, 1995; Tuch et al., 2001; Güllmar et al., 2010]. Also, Connectomes - estimates for the connectivity between different brain areas - can be built from DTI data.

3.1.2.4 Electrical Impedance Tomography

Electrical impedance tomography (EIT) is an image generating technology using electrical impedance measurements [Barber and Brown, 1984; Holder, 2005]. In EIT, information about the distribution of inhomogeneities within some medium (e. g., the human head) is usually gained by applying low currents between some electrodes, and measuring the resulting impedance at another set of electrodes. EIT then uses this information to construct or refine an image of the individual head's conductivity. Compared to CT or MRI the measurements in EIT provide less information for spatial inferences, on the other hand, the application of EIT is much less costly than the acquisition of an MRI or CT scan. In fact, EIT can be performed with conventional EEG and current stimulation hardware commonly available in research and clinical settings. This is done in the following fashion: small harmless currents are injected through pairs of conventional EEG

electrodes into the scalp and the resulting potential is measured on the remaining electrodes all over the head. Successively, all possible pairs of electrodes (e.g. 64 electrodes: $\frac{64\times63}{2}$ combinations) or subsets are investigated.

EIT has recently been employed for head model refinement and verification in neuroscientific contexts [Ferree et al., 2000; Goncalves et al., 2003; Clerc et al., 2005b; Malony et al., 2011; Datta et al., 2013]. However, its use has so far been restricted to estimating tissue conductivities, in particular the skull-to-scalp conductivity ratio. Although the spatial distribution of these conductivities also contains geometric information, EIT has not been used to estimate a shape model of the human head for the purpose of head model individualization. Optimizing shape parameters is therefore one of the key novelties within approaches introduced in *Chapter 5*.

Dedicated EIT devices typically achieve Signal-to-Noise ratios (SNR) in the range of 60dB (1000:1) or even higher [Holder, 2005]. Currents are typically applied at small amplitudes of 100uA and below due to biomedical safety regulations, although higher values are possible (see [Gilad et al., 2007]).

3.2 Biophysical Models

The local distribution of conductivities play a crucial part in the complex electrical interaction between neuronal activity within the brain and it is non-invasive imaging and stimulation. These conductivities are directly connected to the individual head geometry, which can be modeled by more or less realistic head models. Volume conductor models are required both for electromagnetic brain mapping using EEG and MEG as well as for all variants of targeted electrical stimulation and also Electrical Impedance Tomography (EIT). In fact, the same model can be used to model current propagation in both ways from active neural populations to scalp electrodes as well as from scalp injection sites to the entire brain. In the following, a brief overview of the dominant modeling approaches is provided.

3.2.1 Realistic Head Models

Historically, modeling the electric properties of the human head has first been done using spherical models for the brain, skull and skin compartments. The reason was mainly due to the lack of analytic formulas for more complex geometries and the low availability of computational power. However, this coarse approximation to the individual anatomy - where the only parameters are the radii of the spheres - leads to source localization errors of up to 25mm [Yvert et al., 1997]. Improvements were made using spherical harmonics expansions, where the individual head geometry is approximated with equations for which analytic solutions are known [Nolte and Dassios, 2005]. Nowadays these analytic and semi-analytic approaches are almost obsolete due to the fact that practically any

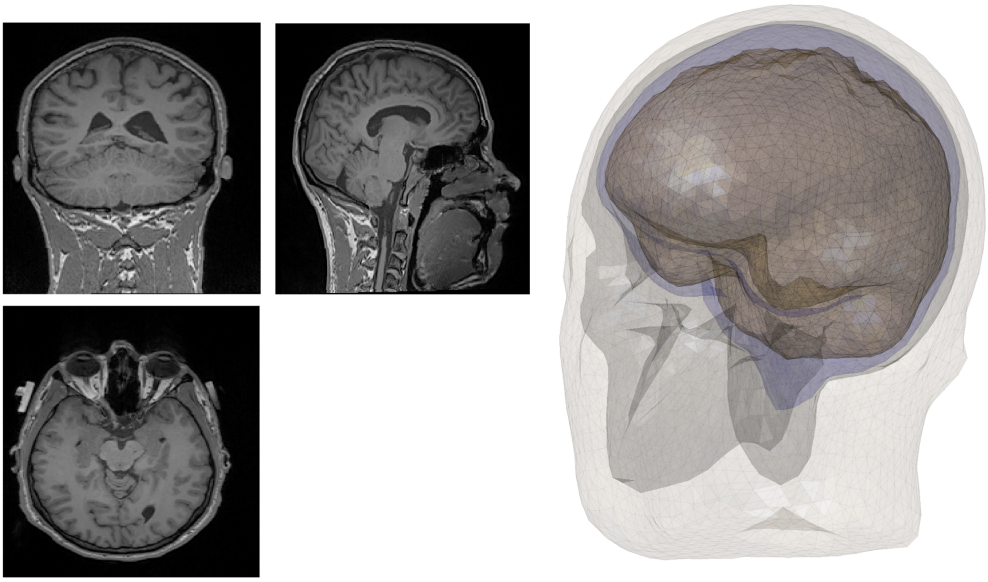

Figure 3.3: A single subject MRI and the corresponding 4-shell BEM head model

standard computer can solve complex numerical models with thousands or even millions of variables within a reasonable amount of time.

Several studies have underlined the importance of realistic volume conductor models of individual heads for obtaining accurate forward model solutions. These studies formulated guidelines for what needs to be included in such models (e. g., [Ferree et al., 2000; Baillet et al., 2001; Wolters et al., 2006; Güllmar et al., 2010; Malony et al., 2011; Dannhauer et al., 2011; Nissinen et al., 2011; Vorwerk et al., 2014]): Not only do the head size and shape vary strongly among people and generally influence the measurements [Nissinen et al., 2011], also the shapes of the inner structures are different. Moreover, tissue conductivities vary between subjects, and can even change within subjects depending on the organism's activities and state. Other main factors varying across people are the thickness, shape and position of the skull, the (inhomogeneous) conductivity of the skull, the size, shape and position of the brain, as well as white matter anisotropy [Wolters et al., 2006; Güllmar et al., 2010; Vorwerk et al., 2014]. In order to obtain a precise head model, the above-mentioned properties need to be estimated on the level of individual subjects.

The 'gold standard' in EEG/MEG studies involving source localization is based on the acquisition of magnetic resonance imaging (MRI) scans of each individual subject. From this, almost all information about the geometrical structure and even most of the different conductivities of the tissues can in principle be inferred. The strongly anisotropic properties of white matter can be estimated using diffusion tensor imaging (DTI) simi-

lar to how the functional Connectomes mentioned in *Section 2.2.2.2* are extracted [Tuch et al., 2001], because both are based on the local orientation of white matter fiber tracts.

3.2.1.1 Extraction of Realistic Anatomy

While an individual geometric model can be obtained up to a very high detail for every subject through MRI and/or computed tomography (CT), these procedures are intensive in cost and labor. Given a subject's individual MRI data, a segmentation of the head into compartments corresponding to different tissue types is obtained using image processing algorithms [Ashburner and Friston, 2005; Huang et al., 2013]. The segmentation information is then used to solve the electromagnetic forward problem. The tissue-type is mostly estimated within each voxel of an MRI by its probability derived from an aligned template. Most approaches involve the Statistical Parametric Mapping toolbox (SPM) [Friston et al., 2007] and use the template extracted from the International Consortium for Brain Mapping (ICBM) average of 152 images (ICBM152) [Mazziotta et al., 2001]. This standard template is not applicable to extract realistic skull geometries as it is limited to tissues within the brain (CSF, white and gray matter). In order to extract realistic geometries, other approaches are necessary. Among a variety of approaches, the new_segment algorithm of the SPM-toolbox [Friston et al., 2007], can incorporate an extended tissue probability map template and can perform this automatically [Huang et al., 2013]. The segmentation includes a non-linear warp between the template and the individual MRIs. Subsequently, an automatic cleanup procedure corrects for errors of the segmentation. After smoothing the mesh with a vertex-based anisotropic flow [Zhang and Haniza, 2006], a BEM head model can be built through, e.g., the openMEEG software [Gramfort et al., 2010; Kybic et al., 2005] (see *Section 3.2.1.3*). This procedure was implemented to run independently and automatically. The final meshes can be examined exemplarily in *Figure 3.3*.

3.2.1.2 Tissue Conductivities

As introduced in *Section 2*, quasi-electrostatic assumptions are usually made upon the properties of the field below 1kHz in the human head. This leads to a linear current/voltage relationship also known as Ohms law. In addition, for electrical methods like EEG or transcranial current stimulation (tCS), the individual tissue conductivity influences the electrical field distributions and thus has an effect on the measurements and stimulations. To simplify the model, the different tissues types are usually approximated by homogeneous conductors with fixed specific conductivities, although they are in general heterogeneous biological tissue with locally and temporally varying conductivity. The homogeneous conductivities can be approximated by standard values [Geddes and Baker, 1967; Gabriel et al., 1996; Oostendorp et al., 2000] or estimated, e.g., by combined MRI or CT with Electrical Impedance Tomography (EIT) [Fuchs et al., 1998;

Ferree et al., 2000; Goncalves et al., 2003; Clerc et al., 2005b; Turovets et al., 2008], by diffusion tensor MRI methods [Tuch et al., 2001], combinations of the two [Sajib et al., 2016] or by data-driven EEG based estimates [Şengül and Baysal, 2012; Hansen et al., 2016; Acar et al., 2016].

Within this thesis, conductivities were approximated by standard values but new approaches to estimating them out of EIT measurements were also developed, see *Section 5.1*. These specific standard conductivities used were $\sigma_1 = 0.465\frac{S}{m}$ for the scalp, $\sigma_2 = 0.01\frac{S}{m}$ for the skull, $\sigma_3 = 1.65\frac{S}{m}$ for the CSF and $\sigma_4 = 0.201\frac{S}{m}$ for the brain. We will see in *Section 5.4.3.3* that the estimated conductivity depends strongly on the model in use and the level of realism in anatomy representation.

3.2.1.3 Numerical Models

After segmentation of the individual 3-dimensional anatomical MRI or CT image into different types of tissues, typically numerical models are applied to estimate the individual electrical properties of the head. This is done either in a volumetric approach on a single voxel level for Finite Element Method (FEM) or Finite Differences Method (FDM) or on discretized sets of boundaries through Boundary Element Method (BEM) based models (see [Vorwerk et al., 2012] for an overview).

The dominant approach in EEG/MEG forward modeling uses a three-shell boundary element model (e. g., [Gramfort et al., 2010]), although finite element models (FEM), finite difference models (FDM) and approaches using semi-analytic expansions into spherical harmonics are also frequently considered [Wolters et al., 2006; Nolte and Dassios, 2005]. Notably, for all these modeling approaches, standard data-analysis pipelines, that require only moderate manual intervention, exist and are often available in open-source toolboxes (e. g., [Oostenveld et al., 2011]). Using these toolboxes, the obtained segmentations might however not always be of acceptable quality. In such case time-consuming manual corrections by trained personnel are necessary.

Whenever the generation of a realistic individual head model is too expensive or simply to time consuming, approximate head models are used. Due to the similarity of the head geometry between people in general, even using simply the head model of another random person works to some extent [Valdés-Hernández et al., 2009]. Another possibility is to implement average head models of databases. Both are to some extent reasonable approximations but in particular very different heads - or outliers in terms of the statistics - are not well represented.

In between average and individual head models, we find approaches that measure the outer head shape and non-linearly warp an MRI or discrete representation of a head to fit this outer shape [Darvas et al., 2006; Acar and Makeig, 2010]. This mostly improves the head model compared to a random person's or most average head models, but, a simple average of co-registered leadfields over subjects was found to perform comparably well as the non-linearly warped template head model [Valdés-Hernández et al., 2009].

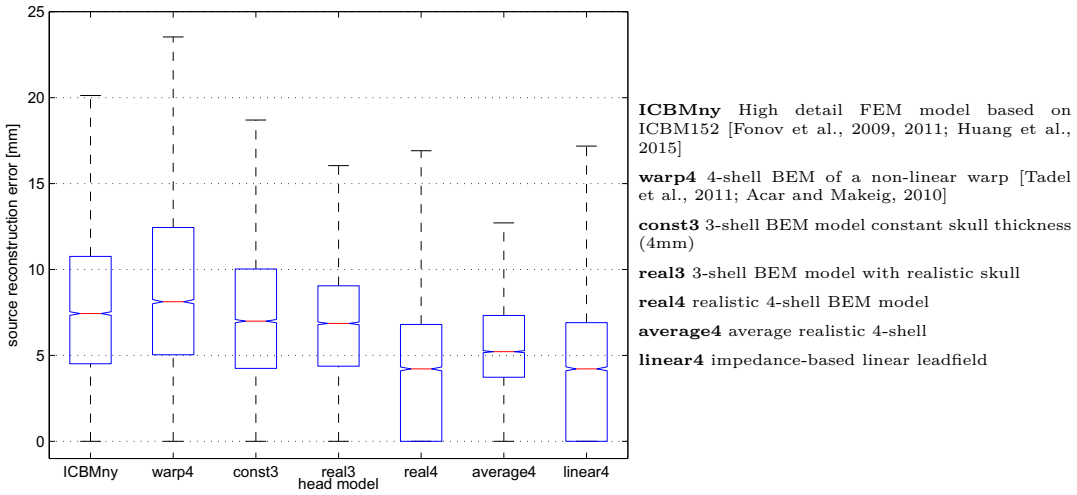

Figure 3.4: The comparison of the quality of different modeling approaches validated on a highly-detailed FEM head model over N=4 subjects in the cortex surface (57 electrodes 10/10 system).

Behind this average over leadfields, a non-linearly warped electrode and source position alignment was applied throughout the database.

Still, all these simpler approaches introduce higher systematic modeling errors than the more sophisticated procedures involving an individual MRI or CT as can be seen in *Figure 3.4*.

Boundary Element Method We will now look at one specific type of head model using the Boundary Element Method (BEM) and the resulting formulas, which will then be used in *Section 5*. In general, if we can divide a space into subspaces Ω_i with boundaries $\delta\Omega_i$ where solutions are known, the use of BEM becomes of interest. As an approximation, this is the case for the human head. Often, BEM is coupled with other methods like FEM or analytic solutions (like the free space dipole solution) in some of the subspaces. In this fashion also anisotropic, inhomogeneous or non-linear solutions for subspaces are possible but the computational advantages mostly vanish over other methods. BEM is most advantageous if we can use simple solutions for most of the compartments. In the following model, harmonic solutions solve all compartments for EIT and for EEG for all except the brain, where an analytic dipole solution is superimposed.

With S_i we define the interfaces between boundary $\delta\Omega_i$ and $\delta\Omega_{i+1}$, so $S_i = \delta\Omega_i \cap \delta\Omega_{i+1}$.

If we assume the conductivity to be piecewise constant in subspaces Ω_i we get a set of coupled Laplace/Poisson equations in the EEG case and only Laplace equations ($f = 0$)

for the EIT case:

$$\begin{aligned} \sigma_i \nabla^2 V &= f && \text{in } \Omega_i, \text{ for all } i = 1, ..., N \\ \nabla^2 V &= 0 && \text{in } \Omega_{N+1} \\ [V]_j = [\sigma \delta_n V]_j &= 0 && \text{on } S_j, \text{ for all } j = 1, .., N \end{aligned}$$

where N is the number of inner compartments of the head and Ω_{N+1} is the space outside the head. Here, conductivity $\sigma_{N+1} = 0$ is assumed and the absence of inner primary currents in EIT leads to $f = 0$ in all subspaces.

The realistic 4-shell BEM extraction in *Figure 3.3* is a procedure created to receive a modeling approach with increased geometric realism compared to common 3-shell BEM and less computational effort than common FEM approaches. It improves common 3 and 4-shell approaches that either strongly simplify the geometries, especially the skull surface shape or otherwise lose their computational advantages over FEM models by having too complex surface discretization. It involves reasonable complexity and lower computation time than common FEM approaches with a relatively high level of anatomic detail compared to common 3 and 4-shell approaches. The reduction in computational effort is advantageous for the size of a head model database we will use. In *Section 5.3*, a low-dimensional representation for the individual head for which BEM models are the better choice is introduced. These can also be used in shape estimations by impedance measurements.

The quality of head models extracted from single subjects is evaluated on a realistic FEM head model in *Figure 3.4*. In this figure, also an average realistic 4-shell BEM and an impedance-based linear headmodel are compared, which we will discuss in detail in *Section 5.4*.

The symmetric 4-shell BEM model used throughout this thesis is based on the set of linear equations defined in [Clerc et al., 2005b]. A symmetric BEM brings certain advantages in terms of accuracies and numerical nature due[Gramfort et al., 2010]. It combines Single- and Double layer approaches, two reciprocal approaches in potential theory [Kybic et al., 2005]. For the linear system equation (compare also *Equation* (2.8))

$$\mathbf{A}_\theta [\mathbf{x}_1, \mathbf{y}_1, \mathbf{x}_2, \mathbf{y}_2, \mathbf{x}_3, \mathbf{y}_3, \mathbf{x}_4]^\top = \mathbf{b} \ , \tag{3.1}$$

where the vectors $\mathbf{x_k} | k \in \{1, \ldots, 4\}$ correspond to the potentials on the interfaces S_k of the different compartments (scalp, skull, CSF and brain), and where the vectors $\mathbf{y_k}$ correspond to its normal currents, we get:

$$\mathbf{A}(\theta) = \begin{bmatrix} (\sigma_1 + \sigma_2)\mathbf{N}_{11} & -2\mathbf{D}_{11}^* & -\sigma_2 \mathbf{N}_{12} & \mathbf{D}_{12}^* & 0 & 0 & 0 \\ -2\mathbf{D}_{11} & \left(\sigma_1^{-1} + \sigma_2^{-1}\right)\mathbf{S}_{11} & \mathbf{D}_{12} & -\sigma_2^{-1}\mathbf{S}_{12} & 0 & 0 & 0 \\ -\sigma_2 \mathbf{N}_{21} & \mathbf{D}_{21}^* & (\sigma_2 + \sigma_3)\mathbf{N}_{22} & -2\mathbf{D}_{22}^* & -\sigma_3 \mathbf{N}_{23} & \mathbf{D}_{23}^* & 0 \\ \mathbf{D}_{21} & -\sigma_2^{-1}\mathbf{S}_{21} & -2\mathbf{D}_{22} & \left(\sigma_2^{-1} + \sigma_3^{-1}\right)\mathbf{S}_{22} & \mathbf{D}_{23} & -\sigma_3^{-1}\mathbf{S}_{23} & 0 \\ 0 & 0 & -\sigma_3 \mathbf{N}_{23} & \mathbf{D}_{32}^* & (\sigma_2 + \sigma_3)\mathbf{N}_{33} & -2\mathbf{D}_{33}^* & -\sigma_4 \mathbf{N}_{34} \\ 0 & 0 & \mathbf{D}_{23} & -\sigma_3^{-1}\mathbf{S}_{32} & -2\mathbf{D}_{33} & \left(\sigma_3^{-1} + \sigma_4^{-1}\right)\mathbf{S}_3 & \mathbf{D}_{34} \\ 0 & 0 & 0 & 0 & -\sigma_3 \mathbf{N}_{34} & \mathbf{D}_{43}^* & \sigma_4 \mathbf{N}_{44} \end{bmatrix}, \tag{3.2}$$

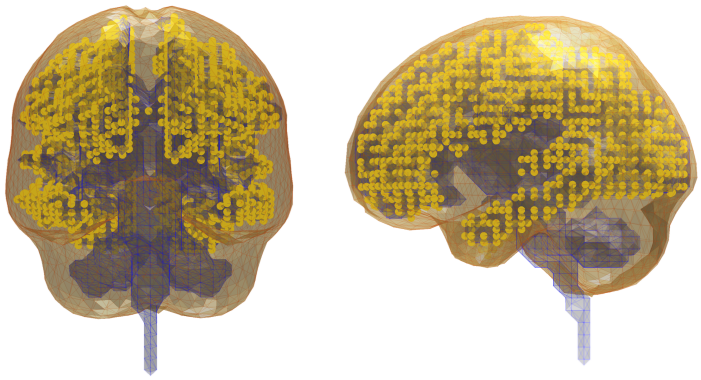

Figure 3.5: Source locations (•) for surface-based analysis on the gray-white matter interface (blue) together with the smoothed cortex surface (light brown).

The matrix elements $\mathbf{N}_{ij}$, $\mathbf{S}_{ij}$, $\mathbf{D}_{ij}$ and $\mathbf{D}^*_{ij}$ are block matrices assembling discretized versions of the integral operators of classical (Newtonian) potential theory between layers i and j that are all based on convolutions with the Green function (*Equation* (2.5)).

For EEG source modeling we have:

$$\mathbf{b} = \begin{bmatrix} \sigma_1 \mathbf{D}^*_{1d}\mathbf{d} \\ \mathbf{S}_{1d}\mathbf{d} \\ 0 \\ 0 \\ 0 \end{bmatrix} \tag{3.3}$$

for discretized source locations in a dipole strength vector or tensor $\mathbf{d}$ which is built from the analytic solution of source density f for the voltage $v = -\frac{1}{\sigma_1} f * G$ by convolution. G is the distance dependent Green function.

For the current injection case (EIT,tCS) we obtain:

$$\mathbf{b} = \left[0, 0, 0, 0, -\mathbf{D}^*_{34}\mathbf{z}, \sigma_4^{-1}\mathbf{S}_{34}\mathbf{z}, \left(-\frac{1}{2}\mathbf{I}_{44} + \mathbf{D}^*_{44}\right)\mathbf{z}\right]^\top . \tag{3.4}$$

Here, $\mathbf{z}$ is the discretization of the normal currents on the scalp and $\mathbf{I}_{ij}$ is an identity operator. It assembles injected currents involving the conductivities σ_i and, within the operators, the shapes of the interfaces.

3.2.1.4 Source Models

The neuronal sources are commonly modeled using an equivalent current dipole on the gray-white interface [Nunez and Srinivasan, 2005; Buzsáki et al., 2012], as it is a common assumption that the main neuronal contributions to EEG come from currents between

the soma (cell bodies) and the afferent synapses located in the dendrites. As outlined in *Section 2.1.1.1*, synchronously active populations of pyramidal cells of the neocortex layer are the main source of neuronal activity measured by EEG. For source simulation, the normal direction of the so generated triangulations can be used to make the contributions more realistic and reduce dimensionality. Still, contributions from all over the brain might be generally found in the EEG. Thus, analysis of source activity can be performed over the whole cortex and source location is then often linked to MRI voxels. (see blue mesh in *Figure 3.5*).

The head models in use include almost exclusively neuronal sources and expect a separate artefact reduction and/or precise experimental procedure avoiding those.

In order to have a comparable set of source locations, the geometric transformations to the segmentation template can be used. This way, sources lie in similar regions of the cortex for different subjects, which is also a basis for average or linear leadfields to work as we will see in *Section 5.4*. In particular, electrode and source locations relative to the surface must meet for the localization to be precise.

3.2.1.5 Gain Matrices: Leadfields for EEG and EIT

After discretization of the geometry, we need to build a model that connects neuronal source activity within the brain to scalp electrode potentials for EEG, scalp electrode currents to scalp potentials for EIT and scalp currents to electric field at source locations for tCS targeting.

Under the assumptions of electrostatics and linearity we can discretize the effects of volume conduction into a set of linear equations like *Equation* (3.1) for all kinds of head models.

Including all the necessary state variables (potentials and/or currents) into a column vector $\mathbf{x}$, we get a linear equation of the form:

$$\mathbf{A}\,\mathbf{x} = \mathbf{b}, \tag{3.5}$$

where $\mathbf{b}$ is a column vector representing the inhomogeneities (the source terms of EIT, tCS or EEG by their boundary conditions on the interfaces). The dimensionalities of the variables depend on the type of head model and the discretizations of the boundaries used.

If $\mathbf{A}$ is regular (as in our case) we can reorder the system in *Equation* (3.5) to:

$$\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}.$$

With an equation of this general form, we can map any known distributions of surface normal currents (like in the EIT and tCS case) or neural source currents (as in EEG source localization) to a potential distribution on the head surface or electrical field

within the brain. If we introduce the interpolation matrix $\mathbf{V}$ of the sensor voltages $\mathbf{v}$ out of the state variables $\mathbf{x}$

$$\mathbf{v} = \mathbf{V}\,\mathbf{x}$$

and similarly divide $\mathbf{b}$ into a forward matrix from sources $\mathbf{B}$ and a vector of scalp currents $\mathbf{i}$ for EIT:

$$\mathbf{b}_1 = \mathbf{B}\,\mathbf{i},$$

or a vector of current dipole moment $\mathbf{s}$ for EEG and another matrix $\mathbf{C}$:

$$\mathbf{b}_2 = \mathbf{C}\,\mathbf{s},$$

we can rephrase the linear system using so called gain matrices. These are often called the leadfield in the EEG and MEG community. We denote them by $\mathbf{K}$ for the EIT gain matrix and $\mathbf{L}$ for the EEG gain (or leadfield) matrix (which can also be a tensor):

$$\mathbf{v} = \mathbf{K}\,\mathbf{i}. \tag{3.6}$$

$$\mathbf{v} = \mathbf{L}\,\mathbf{s}. \tag{3.7}$$

where

$$\mathbf{K} = \mathbf{V}\mathbf{A}^{-1}\mathbf{B}$$
$$\mathbf{L} = \mathbf{V}\mathbf{A}^{-1}\mathbf{C}$$

The $n \times 1$ vector $\mathbf{v}$ stands for the n sensor voltages on the scalp, $\mathbf{i}$ (also $n \times 1$) for the n injected currents on the sensors and $\mathbf{s}$ (dimensionality $m \times 1$) for the current dipole moment of the m supposed source locations within the brain. Accordingly, $\mathbf{K}$ has dimensionality $n \times n$ and $\mathbf{L}$ has $n \times m$. If we want to include all possible source directions (not only the realistic normal to cortex), $\mathbf{L}$ turns into a tensor of size $n \times m \times 3$.

For any linear electrostatic head modeling approach that can link the source currents on the scalp or the source current dipole moments within the brain to the sensor voltages on the electrodes, we can receive the resulting potentials as a superposition of the contributions of the single sources. This can be written in a matrix fashion as the so called gain matrices.

Interestingly, through the principle of reciprocity in electrostatics, $\mathbf{L}$ can also be used to estimate the electrical field strength in the sources imposed by current injection through the electrodes [Wagner et al., 2016].

3.2.2 The Linear Model of EEG

As we have seen in the last section, we can formulate the propagation from sources s to sensors x in terms of a linear relationship as the mixing A is instantaneous. This is often called the linear forward model of EEG [Haufe et al., 2014b]. In the case of EEG,

the number of sources is in general much higher than the number of sensors, making this problem under-determined. Another important factor is that the strongest sources are not neccessarily of neural origin. Therefore, a noise term ϵ is commonly added for all noise that is not taken care of or is not of interest in the source model:

$$x = As + \epsilon \tag{3.8}$$

Within the forward model, the propagation of signals from sources to sensors are stored as patterns in columns of the matrix A. It is in general difficult to measure the source activity directly and the common situation is rather that we have a set of measurements and want to estimate the source activity or localize the most probable sources. This is what is called source reconstruction/separation and source localization, respectively. To this end, the unmixing matrix W has to be estimated, which usually accounts for reducing noise ϵ and estimating orthogonal or independent sources $\hat{s}$:

$$\hat{s} = W^\top x \tag{3.9}$$

As we are operating in a spatial domain with a linear filter, this approach is also termed spatial filtering and can be performed unsupervised (ICA,PCA,..) or in a supervised fashion including prior knowledge about the different points in time (CSP, LDA,...). We will investigate this further in the following chapter.

3.3 Analysis of oscillatory EEG data

We will now focus on the analysis of rhythmic neuronal activity as we face, e.g., in the sensorimotor rhythm (SMR) [Pfurtscheller and Neuper, 2001; Blankertz et al., 2008b, 2010] or estimation of cognitive workload [Gevins et al., 1998; Smith et al., 2001; Holm et al., 2009]. In these, usually mainly a change in amplitude of peaks in the Power-Spectral Density (PSD) related to events in experimental conditions like left/right hand motor imagery or cognitive workload level is investigated. This change is assumed to be due to synchronization/desynchronization effects on the mean field and is thus termed Event Related Synchronization/Desynchronization (ERD/ERS). The related spontaneous changes in background or broadband activity mentioned in *Section 2.2* are to be considered rather in future approaches. Also, the fact that in our simulations SSD suppressing the background activity works best in classifying motor imagery data suggests, that those factors do not play such a crucial role in Brain Computer Interfaces (BCIs) using spontaneous EEG.

Analyzing or classifying rhythmic (or oscillatory) neuronal activity is not a straightforward task, as the signal itself contains a spatially smeared mixture of non-linear amplitude dynamics distributed over different frequency bands. Many of the peaks in the PSD actually covary due to the non-sinusoidal nature of the source signal which is

caused by the non-linear interactions of neurons. Focusing on a single peak, we experience that the amplitude of that signal in time or its power in frequency covaries with task condition. Also, the background activity in shape of the 1/f noise is known to covary with different conditions. Additionally, the different peaks can linearly and non-linearly covary and actually originate from the same source area.

We thus need to transform the time-varying multivariate EEG signal into an amplitude or power variable in order to interpret, relate or classify different conditions. To counteract the effects of spatial smearing it is mostly useful to perform linear demixing before amplitude extraction in order to increase the separability of different sources. EEG sources are expected to mix linear due to the quasi-electrostatic assumption [Häamäläinen et al., 1993] in the usual EEG frequency range below 1kHz and the linear de-mixing is thus optimal for electrophysiological sources in that range. The interaction of the sources, however, is highly non-linear as we have seen in *Section 2.2*. Now, we will revisit and extend common perspectives and approaches to EEG analysis and classification.

3.3.1 Signal Processing

EEG signal processing has two main goals in order to extract the relevant information from a raw signal:

- The reduction of EEG artefacts from non-neural origin
- Linear de-mixing for separating different sources superimposed through volume conduction

These two aims can be, however, combined into one single step or algorithm. There are two domains in which the signals can be preprocessed: the temporal (including frequency) and the spatial. Approaches can work on one or both of the domains simultaneously. In order to decrease the influence of motion and eye artefacts as well as channels with changing or high impedance, rejection of epochs and channels based on variances is among the most common first steps [Blankertz et al., 2008b]. After this, regression based on EOG channels, Principal Component Analysis (PCA) or Independent Component Analysis (ICA) on the whole dataset is commonly performed to identify and exclude artefactual components [Winkler et al., 2011]. PCA and ICA can also be used to separate different sources, while algorithms like CSP performs supervised separation based on class-differences, which can also be neurophysiologically interpreted [Blankertz et al., 2011].

3.3.1.1 Spatial Filtering

We will now focus on approaches based on spatial filters in order to reduce artefacts and or separate relevant sources. The easiest spatial filters are bipolar references or the

(surface) Laplace filter simply taking the neighbors as reference. Current source density (CSD) filtering is based on head models and is related to the surface Laplace filter as the surface Laplacian is the equivalent of the surface source current density [Tenke and Kayser, 2005]. Other techniques usually performed are Blind Source Separation or specific algorithms to identify the different components based on theoretical considerations.

Blind source separation techniques like Principal Component Analysis (PCA) or Independent Component Analysis (ICA) have the drawback of mostly not perfectly unmixing (separating) the artefacts from neuronal components and the need to identify the artefactual components. The imperfect unmixing is mainly caused by the fact that these techniques can only linearly unmix as many sources as channels are present, while additionally often the assumption of identically independently distributed samples (iid) is violated due to the high auto-correlation of EEG.

Algorithms like spatio-spectral decomposition (SSD)[Nikulin et al., 2011] are expected to minimize the effect of artefacts and background activity by prior knowledge of the noise structure. SSD uses the fact that the noise is rather broad-band while the neuronal signals of interest are spectrally bundled in peaks around specific frequencies. An artefact reduction followed by a maximization of class differences is commonly found in the design of the analysis and a consecutive application of SSD followed by CSP has been proposed [Haufe et al., 2014a; Halme and Parkkonen, 2016].

If we want to maximize the covariance between specific conditions of the experiment, Common Spatial Patterns (CSP) is usually the tool of choice [Fukunaga, 1990; Koles, 1991; Ramoser et al., 2000; Blankertz et al., 2008b]. It optimizes spatial filters to maximize the difference between the covariance matrices within the classes. Many different approaches to regularizing CSP in order to overcome some of its weaknesses have been proposed: CSP is prone to the influence of single artefacts and also strongly - although commonly used for dimensionality reduction - to over-fitting to the training set, see [Lotte and Guan, 2011] for an overview. Examples are simply the regularization of the covariance estimation (e.g. by shrinkage) [Lu et al., 2009], invariant CSP (iCSP) [Blankertz et al., 2008a], which uses additional measurements of artefacts for regularization and stationary CSP (sCSP), a version that penalizes non-stationary patterns [Samek et al., 2012]. It is based on similar theoretic grounds as a preprocessing by Stationary Subspace Analysis (SSA) [von Bünau et al., 2009]. SSA finds stationary directions in order to reduce artefacts in the data.

Artefacts are mostly non-stationary, but, additionally, stationary components theoretically generalize better. Problematic is, as we will see in *Section 3.3.2* and *Section 4* that also the relevant sources are most probably non-stationary. It is found in the $\frac{1}{f^{\alpha}}$ structure of the amplitudes of oscillations.

Unsupervised algorithms

PCA Principal Component Analysis (PCA) is a method that detects uncorrelated directions in the covariance of the data [Pearson, 1901]. It is related to an eigenvalue decomposition of the covariance matrix Σ_x of the data x. It can be formulated as an optimization problem for the single filters (eigenvectors):

$$w^* = \arg\max_w \frac{w^\top \Sigma_x w}{w^\top w} \tag{3.10}$$

while this relates to an eigenvalue problem that solves:

$$W^\top \Sigma_x W = D \tag{3.11}$$

$$W^\top W = I \tag{3.12}$$

PCA is often used as an artefact reduction step in EEG as the strongest components are most likely eye and muscle artefacts due to their high variance [Blankertz et al., 2011].

ICA Independent component analysis (ICA) is a common tool for finding statistically independent sources in EEG [Makeig et al., 1996]. In its assumptions, it is similar to PCA but involves approximation of statistical independence. Correlation and independence are related while decorrelation can be seen as one of the most straight-forward approximations of independence. Independence leads to uncorrelated variables but uncorrelatedness does not imply independence as it is based on lower-order statistics only. Hence, PCA is mostly used as a preprocessing step for ICA in order to facilitate convergence of the algorithm. ICA usually adds higher-level statistics in order to improve the robustness to non-Gaussian distributions and non-linear interactions. There are many different ICA algorithms of which fastICA [Bell and Sejnowski, 1995] is one of the most widely used. It uses the Infomax principle related to the entropy in combination with assumption of nonlinear distributions. The Infomax principle is based on maximizing the mutual information between input X and output Y using the entropy of the output $H(Y)$ and the entropy of the output not coming from the input $H(X|Y)$[Bell and Sejnowski, 1995]:

$$I(Y, X) = H(Y) - H(Y|X) \tag{3.13}$$

A non-linear neural network is trained to optimize the target function.

In EEG and similar applications, it often fails at properly de-mixing sources [Lühmann, 2018] as it assumes identically independent distribution of data points which is not given in EEG due to the temporal long-range autocorrelation. Newer approaches consider these facts and lead to better separation of sources [Lühmann, 2018; Adali et al., 2014].

SSD Spatio-Spectral Decomposition (SSD) is a dimensionality reduction technique that is expected to extract neurophysiologically plausible sources [Nikulin et al., 2011]. If we maximize the ratio of the covariance Σ_x of the signal of interest x in a frequency band of interest against the noise in the flanking bands Σ_N in the extracted components, we obtain spatial filters that practically maximize the signal-to-noise ratio of these components. Noise here rather describes everything that is unrelated to the signal of interest: the neural oscillation as a spectral peak in a certain frequency range. The objective function for the optimal spatial filter w^* is

$$w^* = \arg\max_w \frac{w^\top \Sigma_x w}{w^\top \Sigma_N w} \tag{3.14}$$

while this relates to a generalized eigenvalue problem that solves

$$W^\top \Sigma_x W = D \tag{3.15}$$

$$W^\top \Sigma_N W = I \tag{3.16}$$

for the filter matrix W.

SSD is useful as a dimensionality reducing preprocessing step that mainly results in components of neural origin. While the selection of the number of necessary components has been proposed in [Haufe et al., 2014a] by an outlier detection based on the interquartile range, a selection of the 14 best components seems to work comparably. The fact that also broadband activity changes with task condition in the higher frequency band, limits the success of this approach most probably to lower bands like alpha and beta.

Supervised algorithms

CSP The Common-Spatial-Patterns (CSP) algorithm is one of the most successful in the classification of motor imagery based brain-computer interfaces[Blankertz et al., 2008b]. It finds directions in the covariance of the data space that maximize the difference between the classes linked to different tasks (e.g. left vs. right hand movement). There are several different versions of CSP, while the following objective function is used throughout this thesis due to its favorable form in normalizing the output:

$$w^* = \arg\max_w \frac{w^\top (\Sigma_1 - \Sigma_2)\, w}{w^\top (\Sigma_1 + \Sigma_2)\, w} \tag{3.17}$$

This objective function can be solved by generalized eigenvalue decomposition and leads to the following relationships between the filter matrix W, the class-wise covariance matrices Σ_c and the eigenvalues D. For two classes, we get:

$$W^\top (\Sigma_1 - \Sigma_2) W = D \tag{3.18}$$
$$W^\top (\Sigma_1 + \Sigma_2) W = I \tag{3.19}$$

CSP directly maximizes the difference in variance between classes, which is then used logarithmized for classification.

iCSP Invariant Common Spatial Patterns (iCSP) adds covariances of confounding sources Ξ as a regularization term in order to select components that are invariant to these [Blankertz et al., 2008a].

$$w^* = \arg\max_w \frac{w^\top \Sigma_1 w}{w^\top \left[(1-\eta)(\Sigma_1 + \Sigma_2) + \eta \Xi\right] w} \tag{3.20}$$

with the related generalized eigenvalue equation:

$$W^\top \Sigma_1 W = D \tag{3.21}$$
$$W^\top \left[(1-\eta)(\Sigma_1 + \Sigma_2) + \eta \Xi\right] W = I \tag{3.22}$$

sCSP In [Samek et al., 2012], stationary Common Spatial Patterns (sCSP) was proposed. It penalizes non-stationary directions in the data by the term $\overline{\Delta} = \overline{\Delta}_1 + \overline{\Delta}_2$, which consists of the average class-wise difference to the average spatial covariance $\overline{\Delta}_c$ for class $c \in 1, 2$ over trials :

$$w^* = \arg\max_w \frac{w^\top \Sigma_1 w}{w^\top \left(\Sigma_1 + \Sigma_2 + \alpha \overline{\Delta}\right) w} \tag{3.23}$$

The related generalized eigenvalue equation is:

$$W^\top (\Sigma_1 - \Sigma_2) W = D \tag{3.24}$$
$$W^\top \left[\Sigma_1 + \Sigma_2 + \alpha \overline{\Delta}\right] W = I \tag{3.25}$$

The algorithm uses the measure $\overline{\Delta}$ to penalize directions that are non-stationary, which imply large differences to the average trial-wise covariance.

sCSSP Combining the ideas of CSP, sCSP and SSD, sCSSP maximizes the difference between class-wise covariance matrices Σ_1 and Σ_2 while diagonalizing the weighted sum of the noise matrix Σ_N and a non-stationarity constraint $\overline{\Delta}$. The rational behind this is that any two consecutively applied linear spatial filters can be combined into one filter without loss of information and that a combined extraction optimizing all objectives at

once could lead to a selection of better components. The results can be found in section *Section 4.1.* We define the maximization problem as:

$$w^* = \arg\max_w \frac{w^\top (\Sigma_1 - \Sigma_2)\, w}{w^\top \left(\Sigma_N + \alpha\overline{\Delta}\right) w} \tag{3.26}$$

with the connected generalized eigenvalue equation:

$$W^\top (\Sigma_1 - \Sigma_2)\, W = D \tag{3.27}$$

$$W^\top \left(\Sigma_N + \alpha\overline{\Delta}\right) W = I \tag{3.28}$$

LDA Linear Discriminant Analysis (LDA) is an approach that is mostly used for classification in many different contexts including the BCI community [Blankertz et al., 2008b; Congedo et al., 2017]. The assumptions for LDA being the optimal classifier include equal covariances within the classes, Gaussian and known distributions [Friedman, 1989]. Due to the bias in covariance estimation, shrinkage helps to estimate the true covariances [Ledoit and Wolf, 2004]. It nevertheless practically works in all kinds of applications without these being met. It is a generalization to Fisher linear discriminant [Bishop, 2006; Fisher, 1936] and is related to the maximization of the between-class variance S_B against the within-class variance S_N within the Fisher criterion:

$$w^* = \arg\max_w \frac{w^\top S_B w}{w^\top S_w w} = \arg\max_w \frac{w^\top (\mu_1 - \mu_2)}{w^\top (\Sigma_1 + \Sigma_2)\, w} \tag{3.29}$$

We can directly see its connection to CSP from the formula. While not maximizing the filter direction in the covariance between classes it maximizes it in the variance between classes and hence between their means.

3.3.2 Signal Classification

Classification is not only necessary in Brain Computer Interface (BCI) applications but, strictly speaking, every scientific experiment intending to find differences in probability distributions is basically performing a discrimination task.

In classification of spontaneous oscillations in EEG, three predominant approaches are the most widely used:

- Envelope extraction (ERD/ERS) and LDA classification
- Logarithm of variances with LDA
- Riemannian geometric classifiers

Apart from them, there is a variety of approaches including deep learning and other highly involved machine learning based algorithms that, properly validated, do not substantially improve performance [Schirrmeister et al., 2017]. It is debatable, whether the spatial resolution and the relatively low amount of training data are the reason for the success of simple close-to-linear classifiers or the class differences in neural activity are actually that simple.

Usually, the logarithm of the variances is the linearized feature vector of each CSP-channel in use with a consecutive linear classifier. Mostly linear discriminant analysis (LDA) is applied here, see e.g. [Blankertz et al., 2008b].

Already LDA tends to overfit mainly due to bias in the estimation of the covariance matrices which can be partly counteracted by shrinkage [Blankertz et al., 2008b], while this challenge increases with complexity of the classification algorithm. For CSP-filtered data, however, shrinkage is usually not necessary due to the low dimensionality of the data. The main issue seems to be in generalization to unseen data due to the non-stationarities in the EEG.

Another approach uses the theory of Riemannian Manifolds [Congedo et al., 2017; Yger et al., 2016] in order to be less prone to outliers and thus artefacts in the covariance estimations. These approaches are among the best performing state-of-the-art using classification on the tangent space of the geometric mean between the two class covariance matrices. This implies a logarithmization of the distances. The results often lead to the lowest classification errors, while CSP remains a good approximation to these and performs almost as good [Congedo et al., 2017].

In certain settings in more realistic environments simple classification of PSD decompositions can be more successful than CSP and the logarithm of variances. These applications and the relation to increased non-stationarity and the presence of systematic artefacts is discussed in [Miklody et al., 2016, 2017].

Investigating the distribution of single time points over channels from spatial and band-pass filtered EEG data (see *Figure 3.6*), we can see that a Gaussian is the most likely underlying distribution. The images exemplarily show the measured probability density over all data points per class and the difference for one subject. Because of the zero-mean oscillatory structure of the data, most points lie at around zero for both classes. Hence, we need to integrate information over time in order to see a difference between the conditions. This is usually done with either direct or indirect amplitude or (co-)variance extraction in the time domain - the so-called event-related synchronization/desynchronization (ERD/ERS) measures - or with transformation into the frequency or similar domain of short time windows - Wavelet or Fourier approaches. Due to its robustness, the mostly performed approach is the logarithm of the channel-wise variances of CSP-channels. CSP perform a class-wise variance maximization in combination with a separation of sources by decorrelation. Spatially unfiltered data have a much higher covariance among channels.

Another important fact visible in *Figure 3.6* is that the covariances are non-stationary.

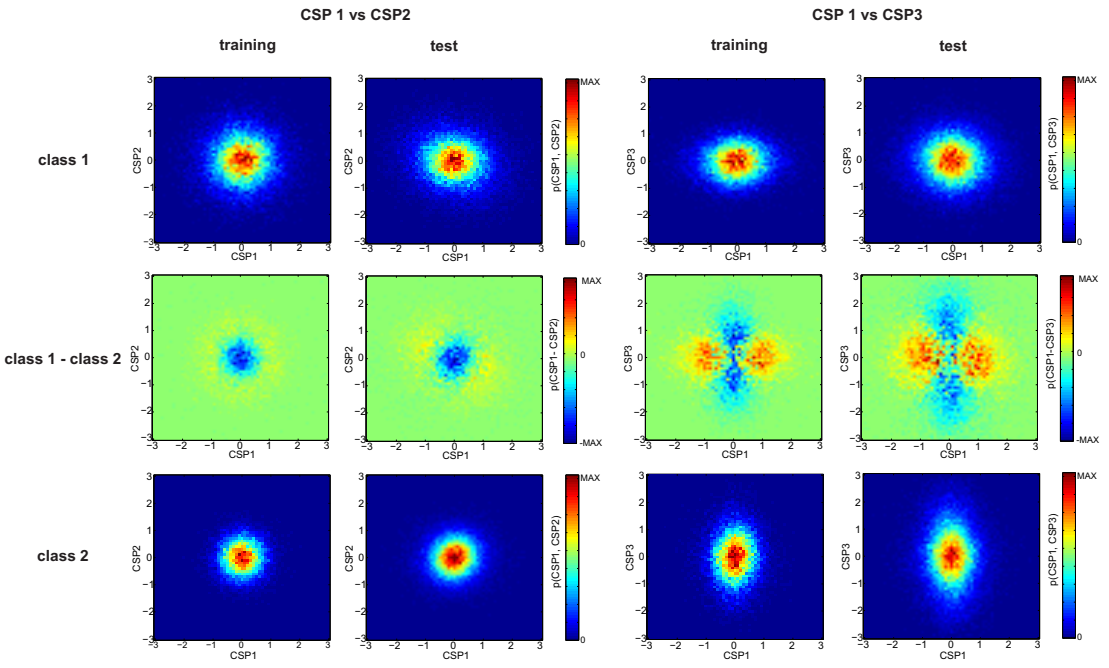

Figure 3.6: The distribution of CSP-spatial and mu band-pass filtered EEG data over different CSP-channels for two classes in a motor-imagery experiment: the information lies in the (co-)variances of the different channels. Another important fact visible is that the covariances are non-stationary in particular in channel 3, the covariance changes from training to test data in magnitude and shape.

The covariance matrices Σ_x of the underlying multi-variate Gaussian distribution change over time for the same condition. This is already present in the training data but as the spatial filtering is optimized on the whole training set, the statistics over the whole set are by definition optimal. Even sub-sets of training data can and usually will, however, be non-optimal in the statistics.

If we investigate the covariance matrices of CSP-channels (*Figure 3.7*), we can see a certain structure of the data caused by the CSP filtering: the channels are de-correlated and the covariance for either of the classes is maximized in each channel. As a result, the variance of a single channel over a certain time window carries the main class information.

Looking at *Figure 3.8* we find that Gaussian distribution is neither given in the variances nor in the logarithm of the variances. Variances of Normal-distributed data is in general χ^2_k-distributed with the degrees of freedom $k = n - 1$ being one less than the number of time points due to the subtraction of the mean. As mentioned, LDA is the optimal classifier for Gaussian-distributed variables with equal covariance in both classes,

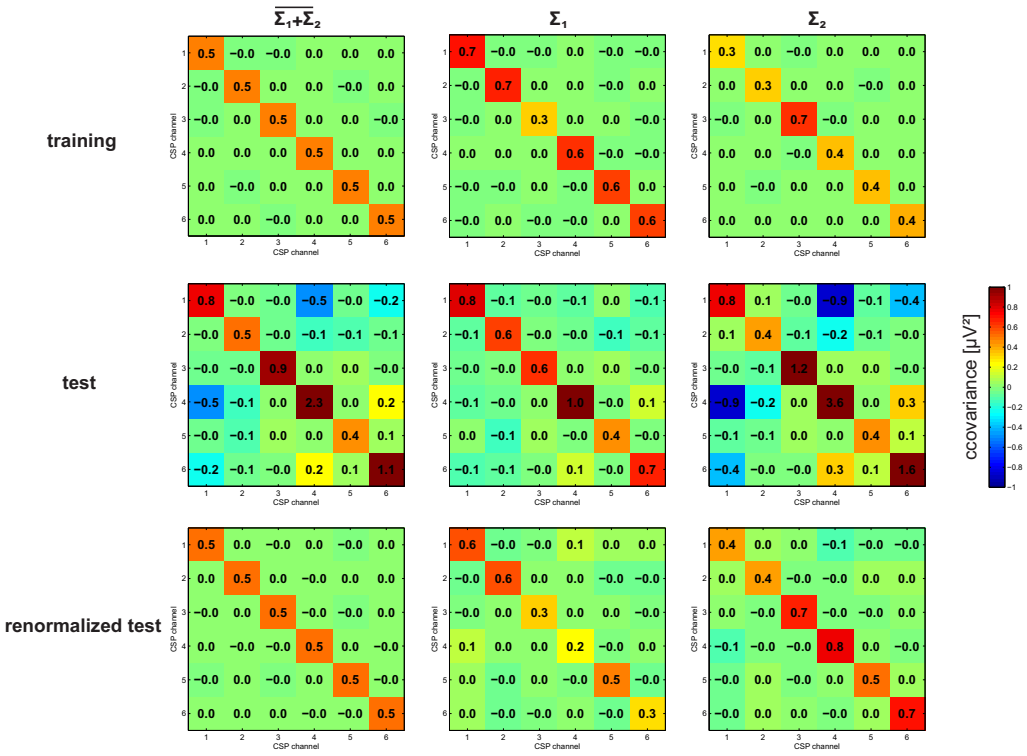

Figure 3.7: The covariance matrices of strongest CSP-channels for training, test and re-normalized test data. A covariance of 0.5 means equal covariance in both classes, 1 full covariance for class 1 and 0 for class 2 data. The test data has been re-normalized for the sum of the class-wise covariance matrices to form an identity matrix. We can see how the magnitude and structure changes for different data with the same experimental conditions. While the first and strongest three channels remain similar in structure, the next three channels are non-transferable to the feedback data as the differences disappear or even flip.

which is neither given. The question arises, why it is still among the best performing algorithms.

The distribution of the variance y for a zero-mean Gaussian-distributed variable x with variance σ_x is (see Appendix A.1):

$$p_y(y) = \left(\frac{k}{\mu_{\sigma_x^2}}\right)^{\frac{k}{2}} \frac{1}{2^{\frac{k}{2}}\Gamma\left(\frac{k}{2}\right)} y^{\frac{k}{2}-1} e^{-\frac{ky}{2\mu_{\sigma_x^2}}} \tag{3.30}$$

where the degree of freedom $k = n - 1$ is one less than the number of samples n on

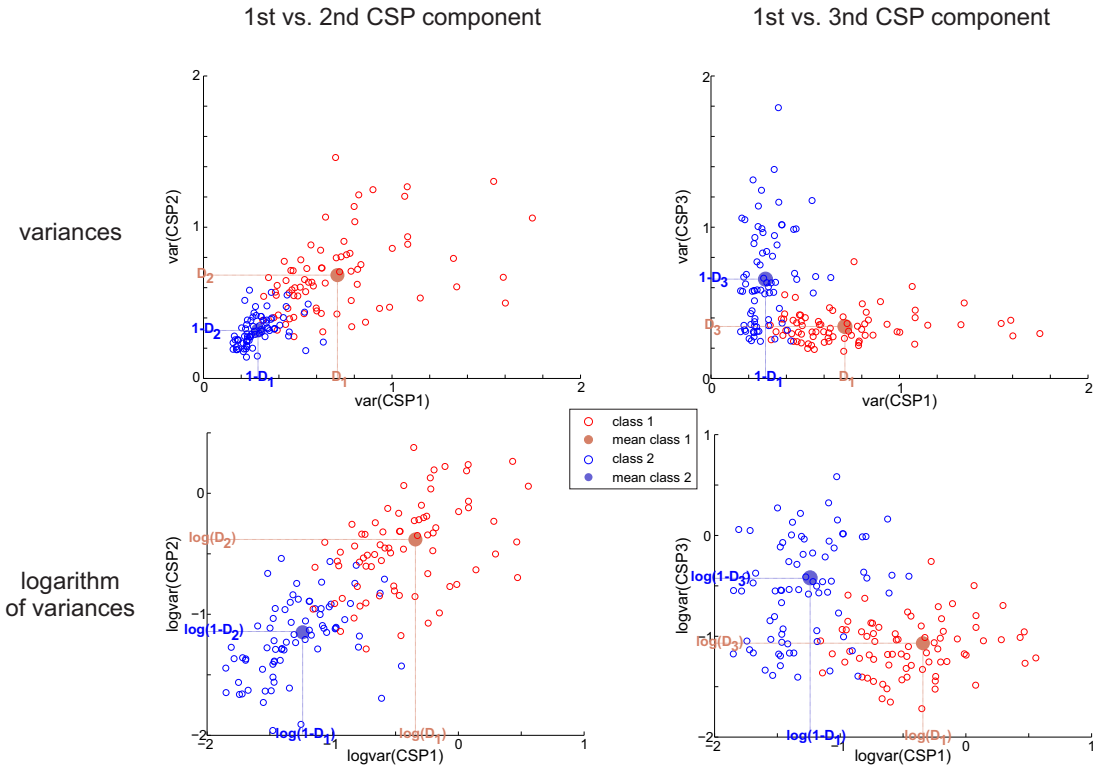

Figure 3.8: The CSP-variances per epoch for one subject in its native and a logarithmic space: the logarithm helps linearizing the data while the results are still not perfectly Gaussian distributed.

which the variances are calculated. The mean μ_y and variance σ_y^2 for CSP-channels with variance d_i are:

$$\mu_y = d_i$$
$$\sigma_y = 2\frac{d_i^2}{k}$$

This holds for a stationary variable and, with larger degrees of freedom k, the variance σ_y quickly decreases relative to the mean, which can be seen in the decrease in coefficient of variation c_V with k:

$$c_V = \frac{\sigma_y}{\mu_y} = \sqrt{\frac{2}{k}}$$

In spontaneous EEG data we often measure the variance of individual epochs with t time points each, which in general leads to a χ^2 distribution with $k = t - 1$ degrees of freedom. As shown in *Figure 3.9*, the non-stationarity of the data, however, leads

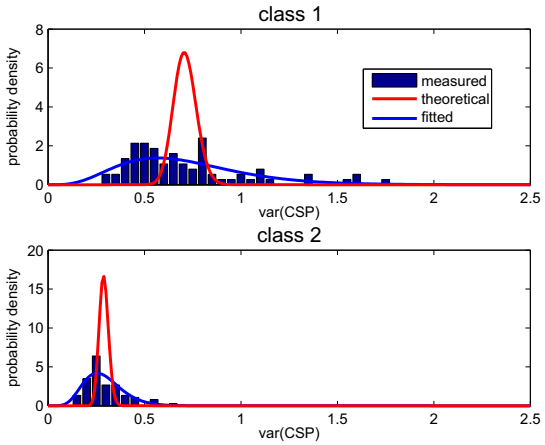

Figure 3.9: The measured and estimated distributions of variances in a CSP-channel for two classes. The non-stationarity in the signal leads to a broader χ^2 distribution than using the degrees of freedom according to the variance theory for stationary signals. For the fitted distribution, the estimated degree of freedom (blue curve) is $\hat{k}^{(1)} = 10.0$ for class 1 and $\hat{k}^{(2)} = 16.6$ for class 2, while for 293 data points of a stationary variable the degrees of freedom would be $k = 292$ (red curve). The histogram is based on one subject and 75 data points per class.

to non-identical distribution of the trials and is even time-dependent due to the $\frac{1}{f^\alpha}$ structure in the amplitudes. This leads to a higher variance of the variance distribution than determined by the degrees of freedom k.

We can fit the distribution to the measurements by estimating a virtual $\hat{k}$ using the measured variance of the variances $\sigma_y^2 = \sigma_{\sigma_x^2}^2$ and their mean $\mu_y = \mu_{\sigma_x^2}$ (see Appendix A.1):

$$\hat{k} = 2 * \frac{\mu_y}{\sigma_y^2} \tag{3.31}$$

This resembles a fit of the generalized χ^2 distribution to the data by the first moments μ_x and σ_x^2. Behind this lies the assumption that the non-stationarity is Gaussian-distributed with a similar structure as the amplitudes themselves. This is justified by the Gaussianity of the distributions of the individual epochs, although a better model might exist. We can now introduce a more general χ^2 distribution which encompasses this effect by using this virtual degree of freedom $\hat{k}$ by *Equation* (3.30). It has to be said again, however that the assumption of identically independent distributed (iid) samples is not valid for EEG with its long-range auto-correlation. The Gaussian distribution of amplitudes seems, however, justified, which leads to a generalized χ^2-distributed variance.

In *Section 4.2*, we will derive optimal classifiers for these kind of distributions, which

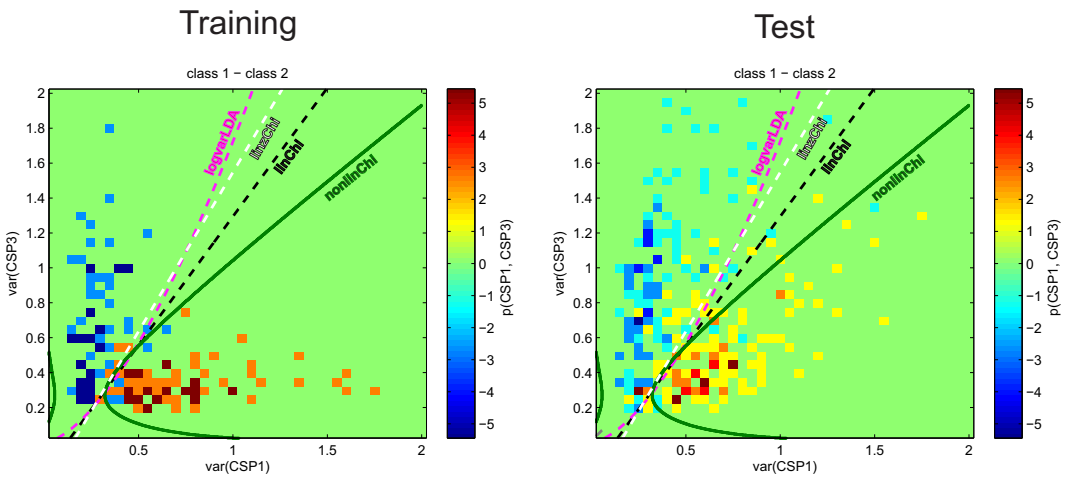

Figure 3.10: CSP-channel 1 vs channel 3 plotted as probability density estimated on the test data and training data. Also, the optimal and log-var LDA classifiers trained on the training data are shown. The different classifiers are logarithm of variances with LDA (logvarLDA), nonlinear optimal χ^2 classifier (nonlinChi), linearized nonlinear optimal χ^2 classifier (linzChi) and the linear optimal χ^2 classifier for equal degrees of freedom per class (linChi).

perform with lower precision but close to LDA on the logarithm of variances for unseen data. In *Figure 3.10*, we can see one reason, why: the estimated distribution changes for the test data and, hence, the classifier is not optimal anymore, while LDA is less affected. The different classifiers are shown together with LDA on the logarithm of variances (logvarLDA) and the distribution of variances and will be further explained in *Section 4.2*.

LDA of the logarithm of variances is less affected by the individual non-stationarities in scaling of the variables per class due to use of the logarithm. The reason is simple and lies in the basic properties of the logarithm. If the change a_i in scale in the n features x_i is different, a linear classifier like LDA $y = \mathbf{w}^\top \mathbf{x} + b$ leads to:

$$y = \sum_{i=1}^{n} (w_i log(a_i x_i)) + b = \sum_{i=1}^{n} (w_i log(x_i)) + \sum_{i=1}^{n} (w_i log(a_i)) + b$$

while for the direct classification in the variances x, we get:

$$y = \sum_{i=1}^{n} (w_i a_i x_i) + b$$

This implies that if the change in scale is on average zero over all channels, it will have no effect for the logarithm but a channel-wise effect for the direct variance classification.

Additionally, the commonly performed bias adaptation will only systematically help in the logarithmic space, as the change in scale is additive. This is also an explanation, why adaptation of the covariance matrix in the LDA classifier usually does not substantially have an effect on the classification of the logarithm of variances [Vidaurre et al., 2011]. Improvements in classification results using pMeans adaptation can be found in *Section 4.1*.

This is similar to the effects of applying Riemannian Geometries, where the special properties of the space lead to an equivalence in distance between source and sensor space [Congedo et al., 2017]. This also implies that non-stationary amplitudes even for spatially non-stationary sources do not affect the distance and, hence a classifier based on it. Also, here, a logarithmization of space is performed which is why logarithm of variances of CSP with its spatial unmixing and approximated scale invariance are a good approximation of some Riemannian classifiers. They just miss the invariance to non-stationarities.

3.4 Chapter Summary

Different methods exist for directly measuring neuroelectric activity. There is non-invasive method with low spatial resolution and broad coverage like EEG on the one hand and invasive with high spatial resolution but only local coverage like intracellular, LFPs and ECoG recordings on the other. MEG is in between, with a higher resolution but a lower coverage than EEG.

While, due to the effects of volume conduction, in particular for EEG spatial filtering is almost inevitable to receive good signals, the non-stationarities of artefacts and the neuronal sources themselves introduce issues with generalization to unseen data. LDA trained on the logarithm of variances of CSP-channels is relatively robust in this sense. Looking at the distributions of the different features reveals also why results are more robust in a logarithmic space: the scale invariance helps to counteract non-stationarities. Another factor is that CSP focuses on the main objective: the separation of classes. Algorithms with the objective of robustifying CSP to artefacts by directly constraining them or non-stationarities do not lead to reliable improvements. This can also be accounted to the non-stationarity of the artefacts and thus an overfitting to these. Only in highly artefact-polluted and non-stationary environments like realistic scenarios, CSP fails and simple approaches based on PSDs can perform superior. Other more recent approaches like Riemannian Geometry completely avoid these issues by the equivalence of source and sensor space in their metric.

Chapter 4

Oscillation-based EEG Analysis

We will analyze in detail the application of some of the most effective tools in spatial filtering of EEG data and combine two of them into one algorithm. By this, we will verify whether it is of advantage to execute those algorithms consecutively, as commonly performed, or we rather combine the extraction into one spatial filter. In the second part of this chapter, we will introduce new classifiers for variance data and evaluate their performance on the same motor imagery dataset. These classifiers are based on the χ^2 distribution and are optimal for stationary variance classification. We will thus investigate the effect of the non-stationarities of EEG on their generalization performance.

4.1 Combining Linear Spatial Filters

We will evaluate the advantages of SSD, CSP and sCSP performed sequentially and combined into one algorithm on the same data. On the one hand, it may be of advantage to receive optimal spatial filters within one single step and on the other hand two sets of linear filters can be mathematically combined into one. We will also investigate the effect of classifier adaptation in the shape of moving average pMean [Vidaurre et al., 2011].

The new approach combining SSD and sCSP, termed stationary Common Spatio-Spectral Patterns (sCSSP), additionally offers the possibility to easily update the spatial filters to non-stationary noise structures. The eigenvalue of the filters directly depicts the signal-to-constrain ratio of the components, which, depending on the hyper-parameter, mixes between signal-to-noise and signal-to-non-stationarity constraints.

4.1.1 Formulation of the Optimization Problem for sCSSP

Combining the ideas of CSP, sCSP and SSD, we designed an algorithm that maximizes the difference between class-wise covariance matrices Σ_1 and Σ_2 while diagonalizing the weighted sum of the noise matrix Σ_N and a non-stationarity constraint $\overline{\Delta}$. We define

the maximization problem as:

$$w^* = \arg\max_w \frac{w^\top (\Sigma_1 - \Sigma_2) w}{w^\top \left(\Sigma_N + \alpha\overline{\Delta}\right) w} \tag{4.1}$$

Here, $\overline{\Delta} = \overline{\Delta}_1 + \overline{\Delta}_2$ with $\overline{\Delta}_c = \frac{1}{K}\sum_{k=1}^{K} \Delta_c^{(k)}$ being the average (positive definite) difference matrix of class c [Samek et al., 2012] (see also *Section 3.3.1.1*). So, $\overline{\Delta}$ corresponds to the penalty term of sCSP connected to non-stationarities and α is the corresponding hyper-parameter. Note that we did not normalize the covariance matrices by their traces as in the original sCSP article in order to keep the meaningful covariance values.

Similar to CSP and other optimization algorithms, we can perform a generalized eigenvalue decomposition in order to obtain the spatial filters connected to the maximal and minimal covariance ratios. If we take all Eigenvectors $\mathbf{w}_i$ into a matrix W, they serve as simultaneous diagonalizers for $(\Sigma_1 - \Sigma_2)$ and $\Sigma_N + \alpha\overline{\Delta}$.

$$W^\top (\Sigma_1 - \Sigma_2) W = D \tag{4.2}$$

$$W^\top \left(\Sigma_N + \alpha\overline{\Delta}\right) W = I \tag{4.3}$$

While the noise and non-stationarity term are whitened in the source space $S = W^\top X$ (see *Equation* (4.3)), the difference between classes is maximized (see *Equation* (4.2)). An absolute eigenvalue above 1 implies the class difference being stronger than the noise and non-stationarity level, which turns out to be the optimal selection criterion for a subset of filters in classification.

4.1.1.1 Online Update of Constraints

As we do not know the class-labels in general, we cannot update the class-wise covariance matrices Σ_1 and Σ_2 in a usual application like online experiments. Hence, we will only update the denominator of the objective function *Equation* (4.1). This also implies that the stationarity constraint $\overline{\Delta}$ must be approximated by a deviation from the pooled covariance over classes in the online update Δ. The constrain terms $C = \Sigma_N + \alpha\overline{\Delta}$ in the denominator at M sample are calculated by:

$$C^{(M)} = \Sigma_N^{(M)} + \alpha\overline{\Delta}^{(M)} = \frac{1}{M}\sum_{i=1}^{M}\left(x_N^{(i)\top} x_N^{(i)} + \alpha\Delta^{(i)}\right) \tag{4.4}$$

for the noise band signal x_N with its i-th sample $x_N^{(i)}$ and $C^{(i)}$ the constrain term at sample i.[1] If we receive a new sample M after $M-1$ samples, then:

$$C^{(M)} = \left(1 - \frac{1}{M}\right) C^{(M-1)} + \frac{1}{M}\left(x_N^{(M)\top} x_N^{(M)} + \alpha\Delta^{(M)}\right) \tag{4.5}$$

[1] Note that the covariance estimation is calculated with the normalization $\frac{1}{M}$ instead of $\frac{1}{M-1}$ for homogeneity with the non-stationarity term and simplicity but can easily be changed.

Updating the filters, the following equation (see *Equation* (4.2)) needs to still hold, only the eigenvalues and vectors can change to new values $\tilde{W} := W^{(M)}$ and $\tilde{D} := D^{(M)}$:

$$\tilde{W}^\top (\Sigma_1 - \Sigma_2) \tilde{W} = \tilde{D} \tag{4.6}$$

The noise covariance and the non-stationarity terms can change due to a covariate shift in the data, leading to a need to adapt the weights to still fulfill the diagonalization of the constrain term (compare *Equation* (4.3)):

$$\tilde{W}^\top C^{(M)} \tilde{W} = I \tag{4.7}$$

This is equivalent to whitening the new noise matrix while still separating the classes in a similar fashion. Our old $W^{(M-1)}$ still fulfills the following equation, helping to simplify the update rule:

$$W^{(M-1)\top} C^{(M-1)} W^{(M-1)} = I \tag{4.8}$$

We can introduce a noise-update transformation V at every time step leading to $\tilde{W} := W^{(M-1)}V$ to avoid the recalculation of the whole covariance matrix (which can be computationally and memory intense).

If we insert $\tilde{W} = W^{(M-1)}V$ into *Equation* (4.7), use *Equation* (4.5), *Equation* (4.8) and define a new matrix A, we yield:

$$\begin{aligned}
\tilde{W}^\top C^{(M)} \tilde{W} &= V^\top W^\top C^{(M)} W V \\
&= V^\top \left[\left(1 - \frac{1}{M}\right) I + \frac{1}{M} W^\top \left(x_N^{(M)\top} x_N^{(M)} + \alpha \Delta^{(M)} \right) W \right] V \\
I &=: V^\top A V
\end{aligned}$$

Equation (4.6) leads to:

$$\tilde{W}^\top (\Sigma_1 - \Sigma_2) \tilde{W} = V^\top W^\top (\Sigma_1 - \Sigma_2) W V = V^\top D V = \tilde{D}$$

Now, we can solve a new generalized eigenvalue decomposition $DV = AV\tilde{D}$ and receive the noise-update transformation V and the new eigenvalues $\tilde{D}$.

We can see that the factor $\frac{1}{M}$ balances between keeping the old covariance matrix and the effect of the new sample. In the case of exact covariance matrices of the so far seen data, this is determined by the number of samples M that were seen before. We can also replace $\frac{1}{M}$ by a parameter λ and make the algorithm adopt to non-stationary covariance matrices over time. This can be of advantage as EEG data is known to be prone to covariate shifts. Also, it has the advantage of not needing to memorize the number of samples. Choosing λ appropriately can additionally help adopt the spatial filters to different data.

Then A turns into:

$$A = (1 - \lambda) I + \lambda W^\top \left(x_N^{(M)\top} x_N^{(M)} + \alpha \Delta^{(M)} \right) W \tag{4.9}$$

4.1.1.2 Pooled Mean Online Adaptation

In order to improve the performance of the algorithms, an adaptation of the LDA classifier bias b is commonly helpful. It was used to adapt sCSSP but also - to compare with fair means - for CSP and the other approaches. The formula that is used is based on unsupervised pMean [Vidaurre et al., 2011] and uses a constant update factor η updating at every single trial (hence building the pooled mean over classes implicitly).

$$b(t+1) = (1-\eta)\, b(t) + \eta w^\top x(t) \tag{4.10}$$

In *Chapter 3.3.2* we have seen that pMeans adaptation with logarithm of variances actually adapts to the non-stationarity in amplitudes of the oscillations.

4.1.1.3 Choice of Components

Approaches like CSP and PCA are commonly used for dimensionality reduction, for which a subset of components has to be chosen. Choosing the components strongly influences the behavior of the algorithms and different approaches can be pursued. Not all extracted filters and components might actually provide useful information and improve classification. For CSP commonly the filters with the highest eigenvalues for any of the classes or a certain number per class is a common choice. The 3-6 largest eigenvalues for either of the classes is a reasonable choice for CSP [Blankertz et al., 2008b]. These procedures are usually justified by cross-validation and/or Skree-plot investigations. We determined the best n components on the calibration set and measured the resulting error on the feedback data.

A similar selection scheme can be applied for sCSSP but, actually, the eigenvalues provide additional information: the class-difference-to-noise-and-non-stationarity ratio. For components with eigenvalues above an absolute value of 1, the actual information content is larger than the noise and non-stationarity and this selection criterion turns out to perform best in our analysis. This is a similar selection procedure as the Kaiser criterion [Kaiser, 1958].

For sCSSP we implemented two approaches:

1. The components were determined based on their eigenvalue: an absolute value above 1 means a higher signal level concerning the class-difference than the constraints (noise, non-stationarity). In addition to that a minimum number was chosen to avoid the selection of no components for subjects with low SNR.

2. Similarly as for CSP, a certain number of the highest eigenvalues were chosen.

For SSD, we simply chose the 15 components based on [Haufe et al., 2014a].

4.1.1.4 Choice of Frequencies

The choice of the frequency bands was performed automatically based on the bi-serial correlation coefficient (r), used signed and squared as a measure of separability between classes. The frequency bands for the signal band were chosen based on the same algorithm described in [Blankertz et al., 2010]. In the analysis involving either SSD or sCSSP, we however had to additionally introduce an estimation of the optimal noise band. This was done based on three approaches:

I The noise band ranging from 75% to 95% of the lower cut-off frequency and from 105% to 125% of the higher cut-off frequency of the signal band.

II The noise band chosen by the difference of the r^2 to 1. So the score for the selection $s = (1 - r^2)$.

III The noise band chosen by the squared difference of the r score to 1. So the score for the selection $s = (1 - |r|)^2$.

Approaches I and II where done with these altered scores and the same peak detection algorithm as the selection of the signal band. This lead to the selection of bands that were not contributing to the class-difference.

Investigating Motor Imagery data, the information is commonly linked to alpha and beta bands. While in the original study, a single band peak was chosen automatically in the range from 8 to 35 Hz, we additionally investigated a selection of two peaks, one in the alpha (8-14Hz) and one in the beta range (15-35Hz), as we do not expect peaks spanning over the two sub-bands. A single clear peak surrounded by non-discriminative noise bands is essential for sCSSP and SSD to work. For comparing equal means, we also validated the split band approach with a multi-band CSP.

4.1.2 Main Experiments: Motor Imagery in a Calibration-Feedback Setting

In this study, BCI-naive participants ($N = 80, 41$ female age 29.9 ± 11.5 years, 4 left-handed) were recorded in an experimental routine consisting of two phases. The experimental design minimized the effect of artefacts like eye-movements or muscle activity on EEG. The details of the study and the original approach using CSP can be found in [Blankertz et al., 2010].

In the first phase, the calibration phase, participants were instructed to imagine right hand, left hand or foot movements. The instruction was given by an arrow displayed behind a fixation cross in the center of a computer screen. These data where then used for determining the time interval, the frequency range and the spatial filters according to CSP and sCSSP and other algorithms. The logarithm of the variance of these band-pass and spatially filtered epochs was then used to train an LDA classifier. The best

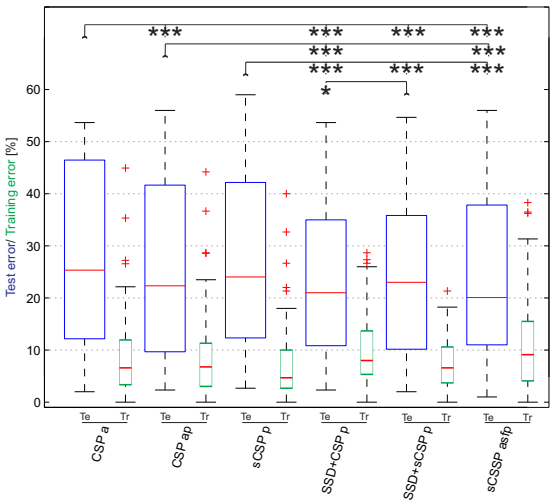

Figure 4.1: **left** Test (Te) and training error (Tr) in the calibration/feedback motor imagery study for different algorithms. sCSSP leads to the lowest median error rates for test data but the highest on the training set. Significances are shown with level (* $\alpha < 0.05$,** $\alpha < 0.01$,*** $\alpha < 0.005$) for the average error being lower in the corresponding column compared to the start of the line indicated by an arrow. Parameters: a - artefact reduction, p - pooled mean adaptation.

performing pair of classes of imagined movements were then chosen for the feedback for each subject.

The second phase was a pure feedback phase, where the participants were instructed again to follow the indication of the target while the cross moved based on a continuous classification in the detected direction to freeze at a final point after 4s.

Different parameter settings were investigated and will be treated in detail in *Section 4.1.4*. These parameters are: a - artefact reduction, p - pooled mean adaptation, f - filter adaptation, s - stationarity constraint.

As can be seen in *Figure 4.1*, sCSSP has the lowest median error of all approaches. All significances mentioned are based on one-sided paired Wilcoxon signed rank test.

The average loss over subjects on classification of sCSSP is significantly lower than of any of the standard CSP approaches. sCSP in our case significantly decreases performance, we could not fully replicate the findings of [Samek et al., 2012] that sCSP improves classification on this same dataset. It has to be said that they used cross-validation to select the parameters for each subject while we used static parameters leading to lower errors. CSP with proper bandpass frequency and interval selection was performing superior to sCSP already and using p-means adaptation even increased the difference. We show sCSP only with adaptation due to the mentioned improvements.

Table 4.1: Mean and median errors for the different approaches.

method	mean	median
CSP	28.9%	25.8%
CSPa	27.9%	25.3%
CSPp	26.9%	23.5%
CSPap	25.4%	22.3%
sCSPp	27.2%	24.0%
SSD+CSPp	23.6%	21.0%
SSD+sCSPp	24.8%	23.0%
sCSSP	23.8%	20.1%

The CSP used for comparison in the original sCSP article seems to perform not up to its capabilities.

Using SSD as preprocessing for CSP significantly improved results while not significantly in combination with sCSP. The difference between SSD+CSP and sCSSP is not significant ($p = 0.62$). Artefact reduction non-significantly decreased average performance for SSD+CSP to 24.2% (not shown), hence artefact reduction was not used in this case.

In figures *Figure 4.2*, the subject-wise classification loss for different sub-approaches of CSP and sCSSP is given. We can see that the largest improvements in both CSP and sCSSP are actually obtained by using pMeans-adaptation. Artefact rejection has little effect. However, using sCSSP significantly improves in all comparisons. Using or not using the stationarity constraint (s) makes up 0.3% improvement. sCSSP already significantly improves the results without prior artefact reduction (compare column sfp). Notably, sCSSP seems to mainly decrease the loss of subjects with CSP loss above $10 - 15\%$ in general.

4.1.3 Validation Experiments: Motor Imagery with Changing Artefact Level

In this experiment, we investigated the performance of the algorithms with further data sets of healthy subjects ((N=28). Additionally, the transfer of the classifier between different artefact levels was evaluated. Likewise, overfitting of the parameters to the data set we optimized them for can be validated.

There were two experimental settings: in one condition, the subjects fixated on the center of the screen, where the motor imagery conditions were shown in form of a letter (R for right hand, L for left hand and F for feet). This minimized eye movements as

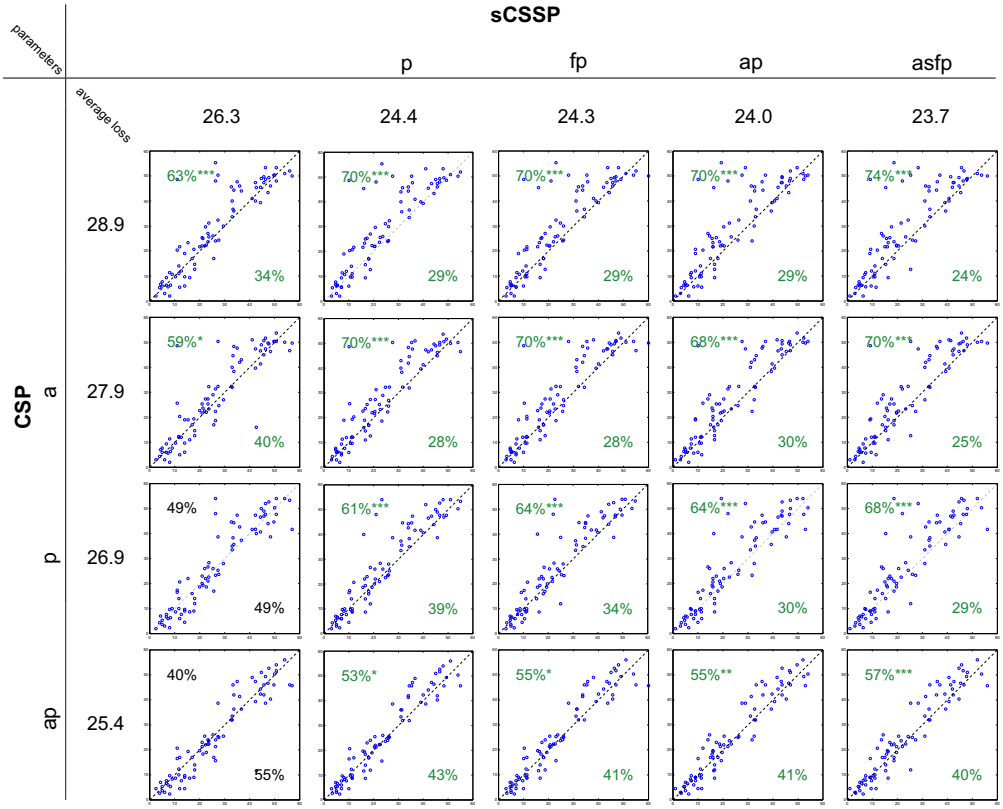

Figure 4.2: A direct subject-wise performance comparison between different CSP and sCSSP approaches. sCSSP significantly improves the results on average (* $\alpha < 0.05$,** $\alpha < 0.01$,*** $\alpha < 0.005$) with the same level of preprocessing and adaptation. The percentages of trials on each side of the equal performance line (dashed) are given (some are exactly on the line). Graphs with green percentages are significantly better for sCSSP and black means no significant differences. Parameters: a - artefact reduction, p - pooled mean adaptation, f - filter adaptation, s - stationarity constraint.

well as neck and jaw muscle artefacts. The same subjects then performed a similar task, where the instruction was constantly moving on the screen in a random fashion, leading to the same task with different artefact level.

Looking at the different classification rates under block-wise cross-validation within the conditions, sCSSP does have a significantly lower average error (18.2%) in the condition with fixation (*fix*) than CSP (20.0%). In the movement condition (*move*), the averages are almost the same (17.7%) and there is no significant difference. The results look similar to those in the other dataset (*Section 4.1.2*) with a performance increase for low

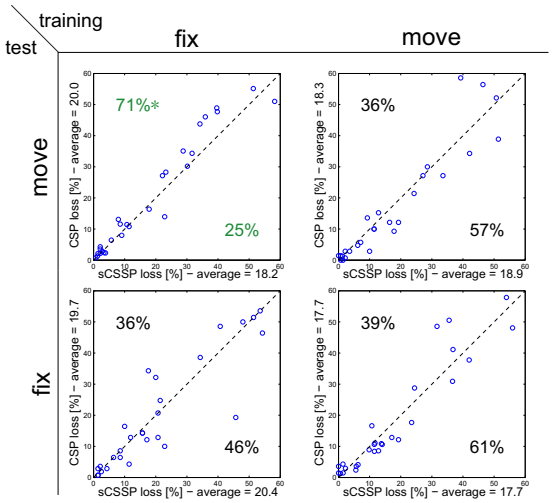

Figure 4.3: Classification results of motor imagery with changing artefact level (*fix*=fixation, *move*=movement) with the optimal parameters of the first dataset ($\alpha = 0.01, \lambda = 1e-4, \eta = 0.05$): an improvement is found by using sCSSP for condition *fix*, while the transfer between *fix* and *move* and the classification on *move* do not change significantly.

performance subjects under CSP and a slight decrease for those with high performance.

If we look at the transfer between the conditions, we find overall rather a slight decrease in performance for sCSSP compared to CSP which is not significant (CSP 19.7%/18.3%, sCSSP 20.4%/18.9%).

Interestingly, the average errors are lower in the condition polluted by eye and muscle artefacts (17.7% CSP/17.7% sCSSP) than in the fixation condition (20.0% CSP/18.2% sCSSP), which has a higher visual alpha level due to the reduced visual input. Additionally, the classification trained on the movement condition also works better in the fixation condition for all approaches shown here (18.3% CSP/18.9% sCSSP). This could be related to the algorithm either choosing components less polluted by artefacts and/or the visual alpha being selected. These differences are not significant.

4.1.4 Effects of Parameter Selection

4.1.4.1 Selection of Components

For CSP, we reached the best results with 3-6 components. We chose 3 for the lowest mean error (25.4% for 3), see the black line in *Figure 4.4*A. The median error for 4 components was slightly lower (22.3% for 3 vs. 21.7% for 4) but the upper quartile was higher (41.7% for 3 vs. 42.8% for 4).

Investigating the blue line in *Figure 4.4*A, we see that sCSSP produces the best results on average by choosing only components with an eigenvalue greater than 1 (25.1%), while setting a minimum of 1 or 2 per band further improves those results (23.8% for 2 components). With more components, the results approach on average those of choosing only a constant number of best components per band based on the eigenvalue (red line). In this case, using only 2 components per band leads to the best results with an average error below any of the CSP approaches (25.1%).

4.1.4.2 Selection of Frequencies

Selecting the noise band frequencies with $s = (1 - |r|)^2$ (approach III in *Section 4.1.1.4* leads to the lowest average error (23.8%), followed by $s = (1 - r^2)$ (approach II, 24.2%). Using a noise band based on relative frequencies around the central peak (I) increases average error (24.8%) non-significantly.

Using multi-band CSP on alpha and beta band leads to no significant difference on the single band values (25.4% for single-band vs. 25.3% for multi-band CSP). The result using sCSSP is different: the multi-band approach improves the results (24.7% for single-band vs. 23.8% for multi-band). This difference is not significant but it is more probable for multi-band sCSSP to improve the results ($p < 0.1$). Also, single-band sCSSP leads to a lower average error than single-band CSP, but this difference is neither significant.

4.1.4.3 Online Spatial Filter Adaptation

The online spatial filter adaptation has only a small effect, in particular when using pMeans adaptation. It improves classification as we can see in *Figure 4.4*B for a range of values around 10^{-4} while the best value is at $1.75 \cdot 10^{-4}$. This lead us to choosing this a a standard value for the other simulations, where filter adaptation was involved. The specific minimum at that value seems related to the dataset it was optimized on, as we could not find similar differences with the validation experiments.

4.1.4.4 Pooled-Mean Adaptation

Pooled-mean adaptation (pMeans) has a significant influence on the loss. The improvement compared to no adaptation for CSP and sCSSP is in the area of 2.5% on the average loss with a smooth convex shape over the tested parameter range. The minimum is at $\eta = 0.05$ for most cases as also reported by [Vidaurre et al., 2011].

4.1.4.5 Stationarity Constraint

Practically, the best results are obtained with low α values which leads to mainly optimizing the class-difference-to-constrain ratio for sCSSP. *Figure 4.4*B reveals an area around

10^{-2} to lead to the lowest average error, while the effect is not smoothly depending on the parameter and interrelated with the settings of other parameters.

4.1.5 Investigations of Scalp Patterns

Investigating the patterns in *Figure 4.5*, we can see clear similarities between the patterns with largest eigenvalues extracted by CSP and those extracted by sCSSP. In general, CSP seems to be more influenced by artefacts in particular if we look at the smallest eigenvalues. In the subjects where CSP performance is superior to sCSSP (tbt,jv,tal), sCSSP has seemingly also caught artefacts in particular in the weakest eigenvalues.

4.1.6 The Effects of Combining Spatial Filters

We have shown, how sCSSP improves classification compared to plain CSP. In a validation set of subjects, we showed how the parameters optimized on the main experiments lead to very similar improvements in a similar setting, but work on the same level as CSP if the artefact levels change. Here, training on a highly (artificially) randomized movement pattern dataset improved the classifier for both CSP and sCSSP, while the difference remains not significant. This effect could be related to a robustification of both algorithms due to the inclusion of non-systematic artefacts in the training data.

Using SSD and CSP separately in a processing chain has a similar effect as sCSSP which is basically a combination of the two algorithms. The advantage of sCSSP is that only one set of spatial filters is learned in one step instead of two filter matrices. The spectra of subjects with lower performance for sCSSP suggest a non-optimal selection of frequency bands in these cases, which could influence sCSSP stronger than SSD+CSP due to the lower total number of components. In SSD, the data is reduced to a higher number of components with lower signal-to-noise ratio before CSP reduces the dimensionality in a second step only based on the class-differences. In sCSSP both decisions are combined and the choice of frequencies and number of components has an interrelated effect. On the other hand, less parameters influence the success of SSD in combination with CSP compared to sCSSP.

In general, a proper frequency selection around an oscillatory peak seems to be the essence of the SSD and sCSSP to work well, similar to the effect of selecting the passband in CSP. A further improved frequency selection procedure could improve results. It is less crucial how wide the noise bands around the central signal peak are but that they do not catch any relevant (here: discriminative) frequencies. If only a small part of the peak is caught by the noise bands, the algorithms minimize the peaks and corresponding components are not selected.

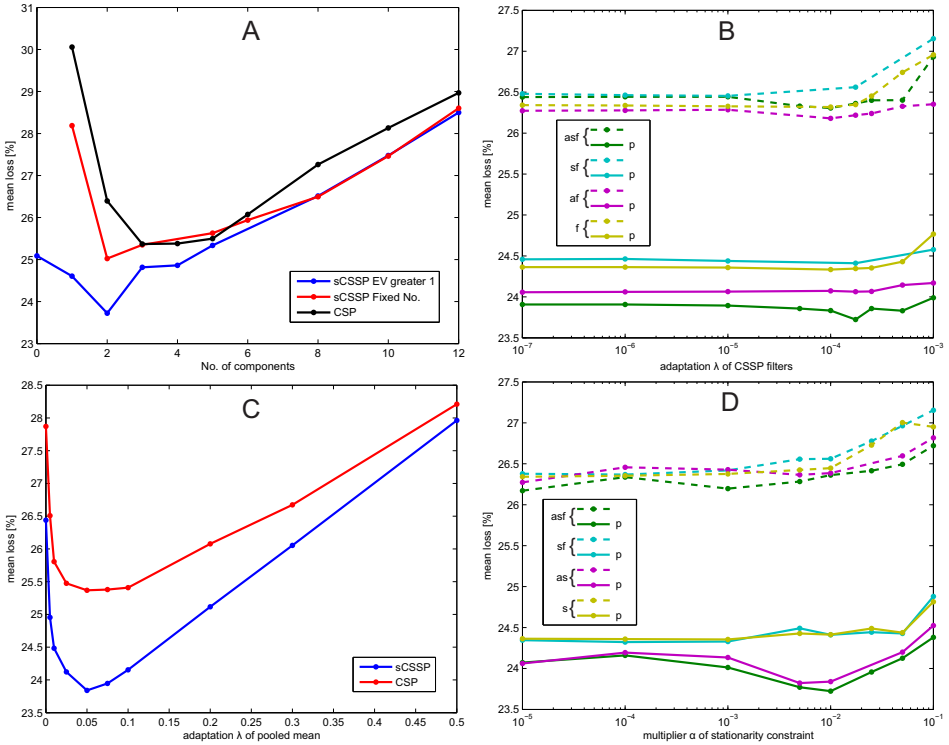

Figure 4.4: The dependence of the average test loss on different parameters: the actual optimal combination of parameters is complex. **A** Number of components per frequency band: an eigenvalue of 1 with an additional minimum of 2 components per channel is the optimum, while a fixed value of three is optimal for fixed selection and CSP. (for the EV criterion the parameter is the minimum number of components per band) **B** The pMeans adaptation has a similarly large effect and a clear optimal value at $\lambda = 0.05$ can be found. **C** The effect of the online adaptation is low and depends on the other parameters. **D** the stationarity constrain has an effect in the optimal setting for the other parameters but little effect else. Parameters: a - artefact reduction, p - pooled mean adaptation, f - filter adaptation, s - stationarity constraint.

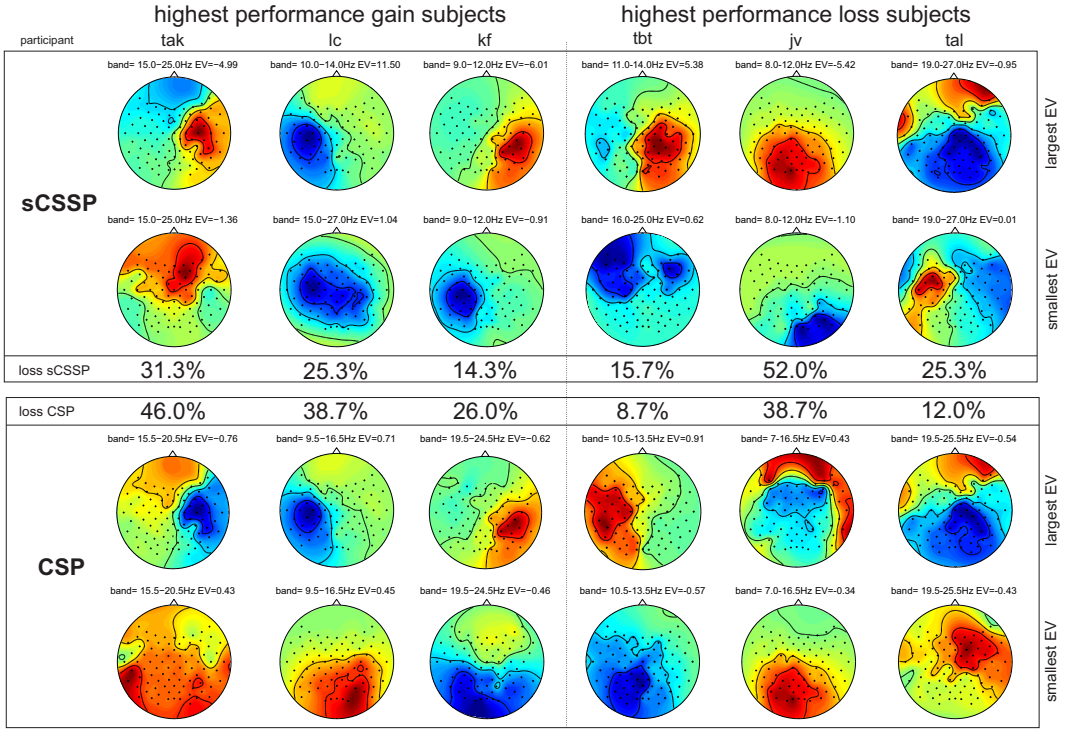

Figure 4.5: Scalp patterns of the three subjects with the highest performance increase between CSP and sCSSP on the left and the highest decrease on the right. We can see a general similarity between CSP and sCSSP patterns, while more CSP patterns seem to be affected by artefacts. One subject with around chance performance (loss> 0.45%) for both algorithms was excluded from this graph. Shown are the largest and smallest Eigenvalues of selected components.

4.2 Distributions of Variance

In the preceding section, one of the most successful classifiers in the context of EEG BCI was implemented: Linear Discriminant Analysis (LDA). LDA is the optimal classifier based on three assumptions: Gaussian distribution of variables, the same covariance within both classes and that the distribution is known. As we have seen in *Chapter 3.3.2*, all assumptions are violated in the case of classification of variance as features as in the previous section. The variance itself follows a χ^2-distribution and the logarithm of it is not Gaussian and neither the same distribution accounts for both classes, see *Figure 3.8*. The variances are non-stationary meaning estimation does dot easily lead to a known distribution. Still, apart from Riemannian Geometric-based classifiers, it is the most successful and most widely used. We now will investigate this and derive another type of classifier that is the optimal classifier for χ^2 distributed data.

4.2.1 Optimal χ^2 Classifiers

As introduced in *Section 3.3.2*, the variance y of Gaussian distributed data x with mean $\mu_x = 0$ and $\sigma_x^2 = d_i$ is distributed according to a generalized χ^2-distribution (see also Appendix A.1):

$$p_y(y) = \left(\frac{k}{\mu_{\sigma_x^2}}\right)^{\frac{k}{2}} \frac{1}{2^{\frac{k}{2}}\Gamma\left(\frac{k}{2}\right)} y^{\frac{k}{2}-1} e^{-\frac{ky}{2\mu_{\sigma_x^2}}}$$

If we have an n-dimensional distribution of variances of n uncorrelated Gaussian variables, we can assume independence as an approximation due to the uncorrelatedness. This leads to the n-dimensional χ^2 distribution being a product of the variance-wise distributions:

$$p(x_1, x_2, ...x_n) = \prod_{i=1}^{n} p_{x_i}(x_i)$$

For unseen data of non-stationary processes, however, this assumption can be violated. In the case of CSP as in *Figure 3.7*, a simple renormalization helps to remain a more stationary and uncorrelated structure. This can be done in online classification in a similar fashion as with pooled covariance adaptation of linear discriminant analysis (LDA) [Vidaurre et al., 2011].

For generalized χ^2 distributed variables, an optimal classifier can be derived (see Appendix A.2). For n stationary independent variables with equal degrees of freedom k_i for all classes within each channel i, the optimal classifier is linear in the variance space:

$$y = \mathbf{w}^\top \mathbf{x} + b \tag{4.11}$$

with $w_i = \frac{k_i}{2}\left(\frac{1}{1-d_i} - \frac{1}{d_i}\right)$ and $b = \sum_{i=1}^{n}\left(\frac{k_i}{2} - 1\right)\log\frac{1-d_i}{d_i}$. In the case of two-class CSP, d_i is the i-th eigenvalue corresponding to the average variance of channel i in class 1, while $1 - d_i$ is the average variance in class 2.

As mentioned, the non-stationarity of the data leads to a much higher variance in the distribution of trial-wise variances. In the case of non-equal (virtual) degrees of freedom $\hat{k}_i^{(1)}$ for class 1 and $\hat{k}_i^{(2)}$ for class 2, the theoretic optimal separating hyperfunction is non-linear (see Appendix A.2):

$$y = f(\mathbf{x}) = \sum_{i=1}^{n}(a_i + b_i \log x_i + c_i x_i) \tag{4.12}$$

which can also be linearized at the 'hotspot' signifying the location on the separating hyper-curve with highest gradient or any other point q in order to receive a linear classifier, see also Appendix A.2). To this and a nonlinear minimization algorithm like the trust region approach [Coleman and Li, 1996] can be used - constrained with the non-linear separating hyperfunction. We will now apply these classifiers on the data from the sCSSP study.

4.2.2 Application to Motor Imagery Data

In *Figure 4.6*, we find the estimated and measured probability distributions between two CSP-channels. Each of them is discriminative for another class in a 2-class paradigm. Also, we can see different separating hyperfunctions of classifiers including:

logvarLDA:	LDA on the logarithm of variances
linChi:	the linear χ^2 classifier for stationary variables (k=292)
nonlinChi:	the non-linear χ^2 classifier for estimated degrees of freedom $\hat{k}$
linzChi:	classifier linearized at the 'hotspot' - the point with largest gradient

We can see that for the displayed CSP-channels, the non-linear classifier is close to an optimal separation for the training data. The test data has changed structure due to the non-stationarities and, hence, the classifier is not optimal anymore, while the shape of the separating hyperfunction of logvarLDA prevents a strong performance decrease. logvarLDA is a linear classifier in the logarithmic space, leading to a hyperplane. In the variance itself it is a hyperfunction of exponential shape.

CSP either overfitted to the training data, the non-stationarities shifted the data or both. Actually thinking about how precise the variance measurements should be according to the underlying χ^2_k distributions with k=292 (compare *Figure 3.9*), the data is obviously non-stationary while CSP is known to overfit[Blankertz et al., 2008b], so both is most probably true.

In *Figure 4.9* we can see the different separating hyperfunctions in logarithmic space. LDA is linear, while the others are non-linear. We have seen in *Section 3.3.2* that the logarithmization in LDA classifications of the logarithm of variances transforms single

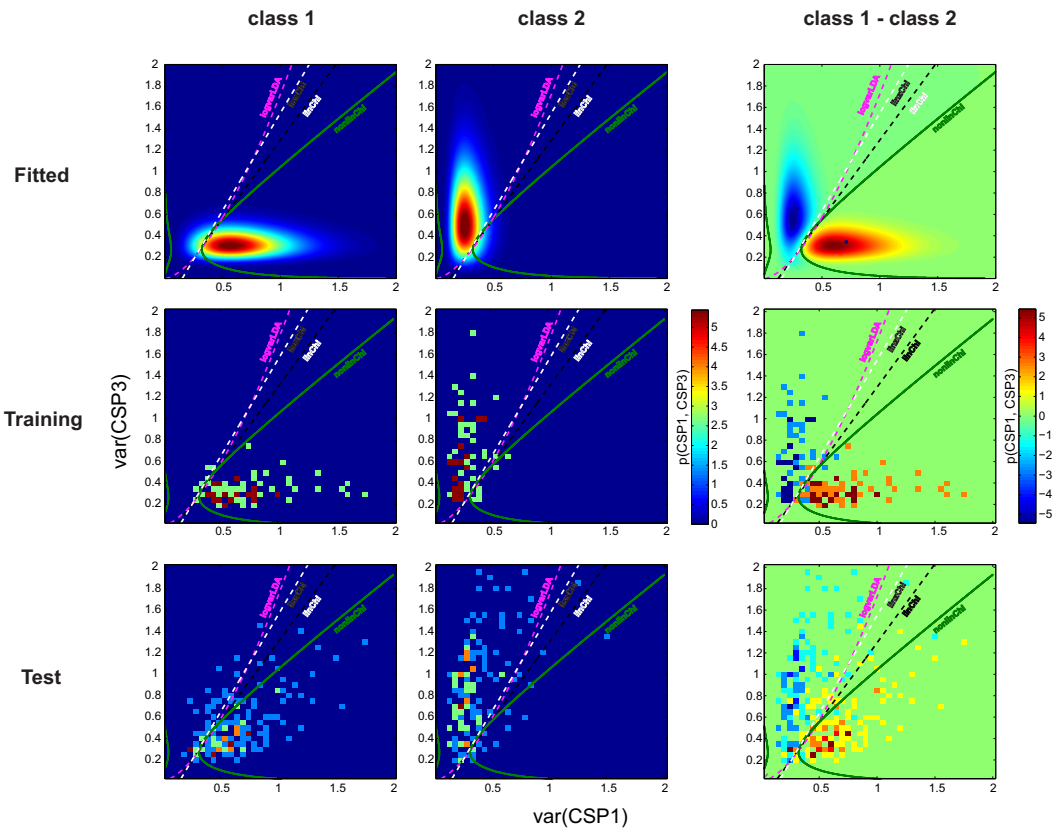

Figure 4.6: CSP-channel 1 vs channel 3 plotted as the fitted chi^2-distribution of training and estimated probability density of training and test data. Also, the optimal and log-var LDA classifiers are shown trained on the training data only. The different classifiers are logarithm of variances with LDA (logvarLDA), nonlinear optimal χ^2 classifier (nonlinChi), linearized nonlinear optimal χ^2 classifier (linzChi) and the linear optimal χ^2 classifier for equal degrees of freedom per class (linChi).

channel amplitude changes into an additive component. If this is on average zero over all channels, the classifier output is not affected. By this, the robustness of this approach can be explained.

Observing the results over the database described in *Section 4.1.2* reveals a large incongruence between training and test errors. The χ^2 based classifiers have obviously been overfitted as their training error is on average lower than for LDA (not significant) but higher on the test data (significant $\alpha \leq 0.005$). The classifiers themselves are hyperplanes in the variance space so probably one of the lowest possible complexities. The problem with generalization is thus rather found in the non-stationarity of the data.

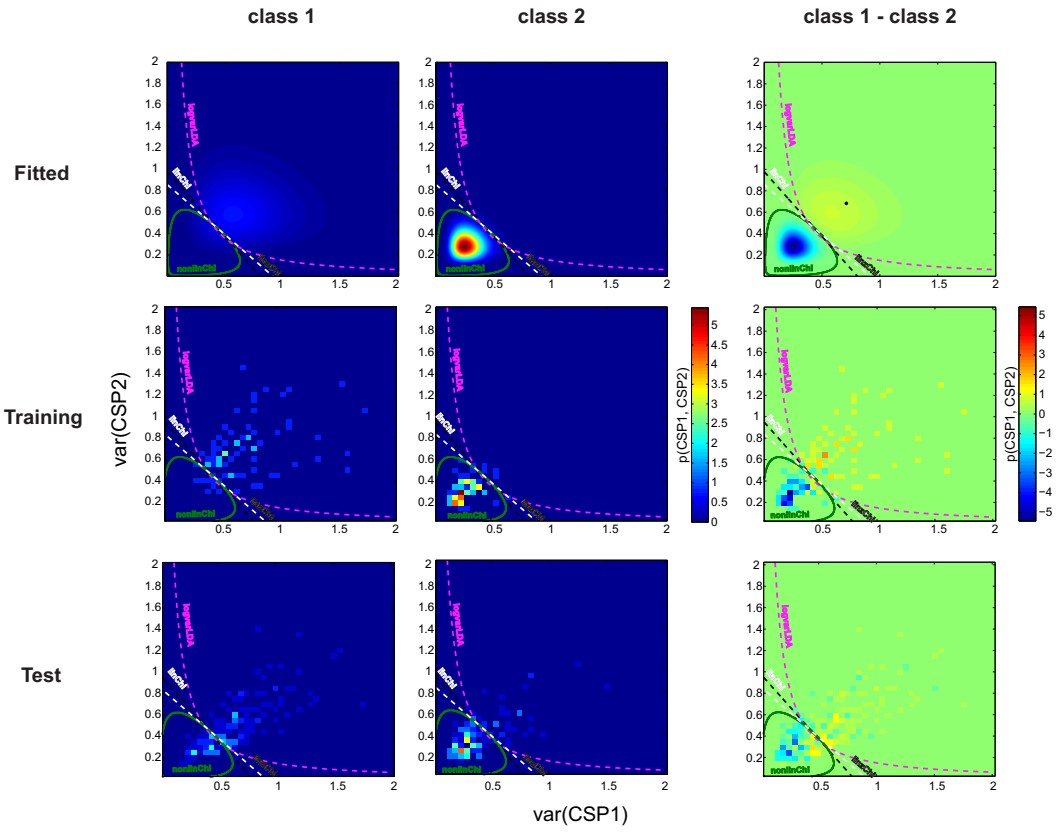

Figure 4.7: CSP-channel 1 vs channel 2 plotted as the fitted chi^2-distribution of training and estimated probability density of training and test data. Also, the optimal and log-var LDA classifiers are shown trained on the training data only. The different classifiers are logarithm of variances with LDA (logvarLDA), nonlinear optimal χ^2 classifier (nonlinChi), linearized nonlinear optimal χ^2 classifier (linzChi) and the linear optimal χ^2 classifier for equal degrees of freedom per class (linChi).

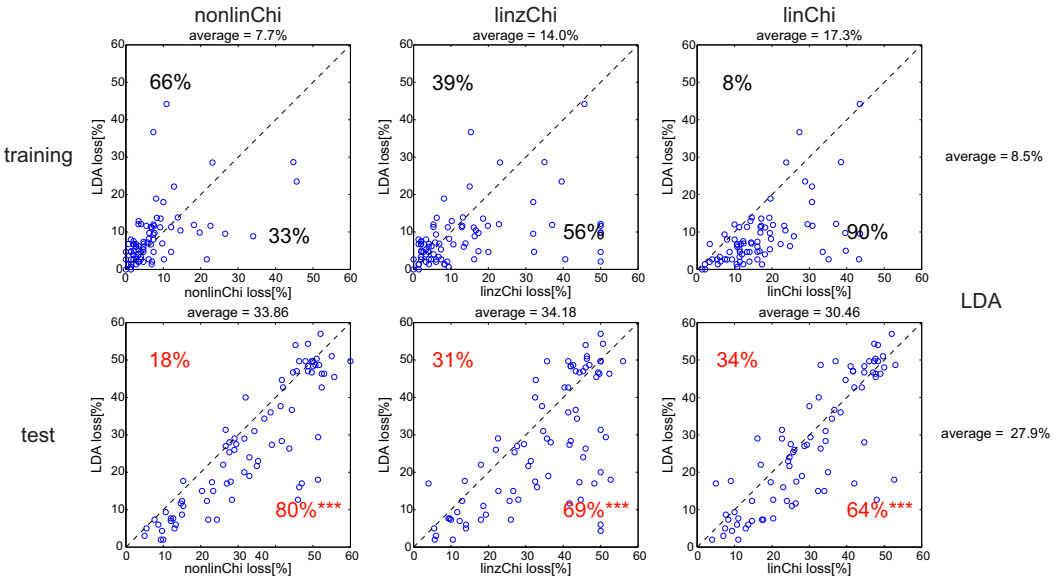

Figure 4.8: Test (first row) and Training (second row) losses of different sub-approaches compared to LDA. We can see lowest training error for the nonlinChi while it leads to the highest test error. The other approaches are in between LDA and nonlinChi. The reason is most probably the non-stationary structure of the signal against which LDA is more robust.

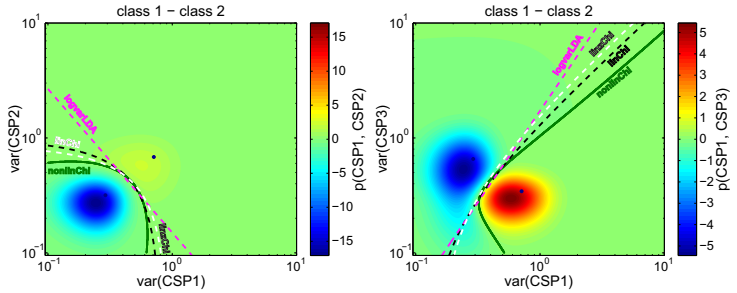

Figure 4.9: CSP-channel 1 vs channel 2/3 plotted as estimated probability density in the logarithmic space together with the separating hyperfunctions. logvarLDA defines a hyperplane here, while the other approaches are non-linear hyperfunctions.

4.3 Chapter Summary

The novel sCSSP algorithm is successful in improving results. Combining the algorithms leads to similar results as applying SSD and CSP sequentially. Said differently, combining linear spatial filters increases complexity but does not have an advantage on classification for the algorithms considered here. It seems that combining too many objectives into one optimization introduces many inter-dependencies that are hard to tackle. sCSSP does improve performance if appropriately trained on data with similar statistics particularly in those of artefacts and noise. Hence, in this case it is well justified to use SSD and CSP as consecutive spatial filters and this drastically increases performance. Also, most probably due to the non-stationarity of the signal, constraining the influence of non-stationarities onto the spatial filters as with sCSP does not necessarily improve performance on average.

A large effect on classification is found in pMeans adaptation, which adapts the classifiers to unseen data. For logarithm of variances classification the adaptation of the bias directly relates to the non-stationarity of the amplitude over time.

Variances are in general χ^2-distributed for which optimal classifiers were derived within this chapter. These lead to lower training and higher test errors on motor-imagery data compared to logarithm of variances LDA classification but reasonable performance on average. The decreased quality in generalization can be accounted to the non-stationarity of the variance distributions that leads to an overfitting to the training data. The logarithm of variances is less prone to amplitude changes due to its transfer of this scaling into an additive component.

Chapter 5

Impedance Based Head Models

Inhomogeneity estimation in conductivity is the original purpose of Electrical Impedance Tomography (EIT). Therefore, a large body of literature exists [Munck et al., 1999; Ferree et al., 2000; Goncalves et al., 2003; Clerc et al., 2005b; Turovets et al., 2008; Malony et al., 2011]. The low technical burden in use makes it appealing for application, while its precision remains low due to the underlying ill-posed inverse problem. We will now look at different approaches that use EIT nevertheless by utilizing appropriate low complexity in the models. The aim is to use EIT for individual head model estimation on the fly, while preparing for an EEG experiment. We will develop algorithms for proper conductivity estimations, electrode position estimation, dimensionality reduction of head models and a linear approach to leadfield individualization.

5.1 Conductivity Estimations with Bounded EIT

For realistic head modeling in general and satisfactory EEG source localization and tCS targeting in particular, a good volume conduction model is crucial. A commonly overseen parameter are the individual absolute conductivities of the different tissues of the subjects' heads. Mostly, rather theoretical, generic values are used and if estimated, these estimations are limited to in-vitro or to the scalp-skull conductivity ratio (SSCR) through bounded Electrical Impedance Tomography (bEIT) [Ferree et al., 2000; Goncalves et al., 2003; Clerc et al., 2005a; Turovets et al., 2008; Fernández-Corazza et al., 2016]. [Clerc et al., 2005a] conclude that the sensitivity of EIT is insufficient for estimating inner conductivities, while geometric model errors showed up to be the actual problem more recently for the SSCR [Fernández-Corazza et al., 2016]. We will see that using a more appropriate optimization algorithm than [Clerc et al., 2005a] leads to the successful and highly precise estimation of all 4 absolute parameters in simulations using 4-shell BEMs while the generalization to more complex data is limited, most probably due to the systematic errors in the model and not the amplitudes or the signal-to-noise ratio (SNR).

We define the conductivity as parameters and tune them to optimally link our measurements to the model. In this manner conductivities have already been estimated by

EIT before, using 3-sphere head models [Goncalves et al., 2000], a 4-sphere model [Ferree et al., 2000], 3-compartment Boundary Element Methods (BEM) [Goncalves et al., 2003; Clerc et al., 2005b] and also using Finite Differences Methods (FDM) [Turovets et al., 2008] and Finite Elements Methods (FEM) [Malony et al., 2011]. [Malony et al., 2011] have even successfully used EIT for estimating the skull inhomogeneity and white matter anisotropy values for a known geometry.

Most authors use either first order optimization like gradient descent [Clerc et al., 2005b] or stochastic optimization like the downhill simplex algorithm [Ferree et al., 2000] or simulated annealing [Malony et al., 2011]. The usage of higher order algorithms was somewhat limited to Gauss-Newton [Nissinen et al., 2015] or quasi-Newton like Broyden–Fletcher–Goldfarb–Shanno algorithm [Horesh et al., 2004], which only partly account for the interdependencies of parameters. [Fernández-Corazza et al., 2013] introduced the classical Newton algorithm involving an FEM head model.

We will now develop and apply a Newton Algorithm to a 4-shell BEM head model. This involves the derivation and implementation of analytical Gradient and Hessian calculations for this type of model and the use of an unregularized modified Newton algorithm to estimate the absolute conductivities within each homogeneous subspace of the head. We can show the success in simulations while a generalization to more complex models (FEM) remains problematic. However, this seems due to the compensation of systematic model errors rather than a problem of the optimization algorithm.

MRI ($1\,mm^3$ isotropic resolution, T1-weighted) of four healthy individuals, all Caucasian male, age range 27–45 at a magnetic field of 3 T (see [Huang et al., 2013]) were acquired. Sets of 92 electrodes spanning down to the neck will be used. For this model, specific conductivities are $\sigma_1 = 0.465\frac{S}{m}$ for the scalp, $\sigma_2 = 0.01\frac{S}{m}$ for the skull, $\sigma_3 = 1.65\frac{S}{m}$ for the CSF $\sigma_4 = 0.126\frac{S}{m}$ for the gray matter and $\sigma_4 = 0.276\frac{S}{m}$ for the white matter.

From these MRIs, A BEM head model can be built involving the openMEEG software [Gramfort et al., 2010]. Two sets of simulations will be conducted in order to investigate the success of bEIT for absolute conductivity estimation in BEM headmodels. Starting values are chosen with a variance of 30% around the true values ($\sigma_1 = 0.465\frac{S}{m}$ for the scalp, $\sigma_2 = 0.01\frac{S}{m}$ for the skull, $\sigma_3 = 1.65\frac{S}{m}$ for the CSF and $\sigma_4 = 0.201\frac{S}{m}$ for the brain.).

First, a corresponding BEM head model with additional Gaussian white noise will be used for the simulation of EIT measurements, after which the same procedure is repeated on a differently generated more realistic FEM headmodel involving 6 tissue types from manual segmentation [Huang et al., 2013]. By this, we will investigate the external validity in form of the generalization to more complex models.

5.1.1 Error Function and Conductivity Derivatives

As it is common in the EIT literature [Clerc et al., 2005b; Goncalves et al., 2003; Ferree et al., 2000; Malony et al., 2011], we can use the sum of squares of the difference between modeled and measured potentials as the cost function to be minimized. The cost function is defined as

$$E(\theta) = \frac{1}{2} \left\| \mathbf{v} - \hat{\mathbf{v}}\left(\theta\right) \right\|_2^2 = \frac{1}{2} \sum_{k=1}^{n} \left(v_k - \hat{v}_k\left(\theta\right) \right)^2 \ , \tag{5.1}$$

where considering the linear head model depending on the parameter(s) θ is:

$$\mathbf{A}\left(\theta\right) \mathbf{x} = \mathbf{b}\left(\theta\right)$$

and its scalp voltage estimate as a linear interpolation:

$$\hat{\mathbf{v}}\left(\theta\right) = \mathbf{W}\hat{\mathbf{x}}\left(\theta\right)$$

leads to

$$\hat{\mathbf{v}}\left(\theta\right) = \mathbf{W}\mathbf{A}\left(\theta\right)^{-1} \mathbf{b}\left(\theta\right) \tag{5.2}$$

For the 4-shell BEM model considered here, derivatives of the cost function *Equation* (5.1) w. r. t. conductivity parameters can be obtained analytically. The first derivative w. r. t. to any model parameter θ_i is according to the chain rule given by

$$\frac{dE}{d\theta_i} = -\sum_{k=1}^{n} \left(v_k - \hat{v}_k\left(\theta\right) \right) \frac{d\hat{v}_k\left(\theta\right)}{d\theta_i} \ . \tag{5.3}$$

and the second derivative:

$$\frac{d^2E}{d\theta_i\theta_j} = \sum_{k=1}^{n} \frac{d\hat{v}_k\left(\theta\right)}{d\theta_i} \frac{d\hat{v}_k\left(\theta\right)}{d\theta_j} - \left(v_k - \hat{v}_k\left(\theta\right) \right) \frac{d^2\hat{v}_k\left(\theta\right)}{d\theta_i d\theta_j} \ . \tag{5.4}$$

The variables modeled by the BEM equations are $\left[\mathbf{x}_1, \mathbf{y}_1, \mathbf{x}_2, \mathbf{y}_2, \mathbf{x}_3, \mathbf{y}_3, \mathbf{x}_4\right]^\top = \mathbf{A}^{-1}\mathbf{b}$, where the model scalp potentials $\hat{\mathbf{v}}\left(\theta\right)$, which are of interest here, are contained in $\mathbf{x}_4$. The derivatives $\frac{d\mathbf{A}}{d\sigma_i}$ and $\frac{d\mathbf{b}}{d\sigma_i}$ can easily be found, since $\mathbf{A}$ contains the conductivities σ_i only as block-wise scalar multiplicands (see *Equation* (3.2)). The corresponding equations for the 4-shell BEM head model can be found in Appendix B.

The modified Newton algorithm for the vector of conductivities σ is:

$$\sigma(t+1) = \sigma(t) - \eta \frac{\nabla E(\sigma(t))}{\nabla^2 E(\sigma(t))}; \tag{5.5}$$

where η is a step size that is one unless an error increase takes place. If so, the step size is decreased until the error decreases or it reaches a small value (2^{-10}).

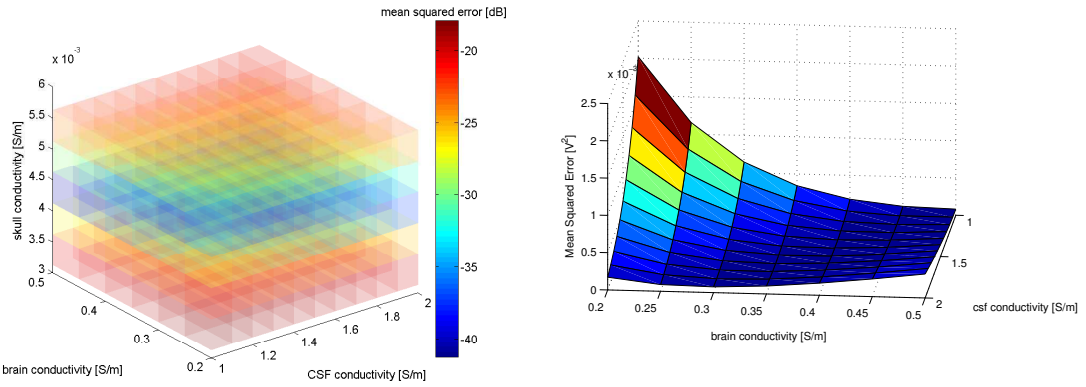

Figure 5.1: The shape of the error depending on conductivity: the error is smooth across parameters, while the sensitivity to either of them strongly varies. It mainly depends on the conductivity of the skull, while dependence on inner structure conductivities (CSF,brain) is rather low and strongly interdependent between the parameters.

5.1.2 Feasibility of Conductivity Estimations

As can be seen in *Figure 5.1*, the optimization problem is smooth also across parameters. There is a very narrow error valley in particular for the skull conductivity. In the 3-dimensional representation on the left, we can see that the mean squared error (MSE) mainly depends on the conductivity of the skull, while is less sensitive to the CSF and brain conductivity (note the dB-scale). On the right, we can see the dependence of the MSE in a linear scale on the conductivities of brain and CSF. Here, the dependency is almost equal on both factors as we can see a valley of low error that leads to almost equal error with proportional scaling of both parameters. This leads to the expectation that this narrow error valley is sub-optimal for first order optimizations that tend to oscillate around this valley due to the higher gradient normal to it than along it.

Gradient descent with individual step sizes for the different tissues can account only partly for the strong difference in derivative towards the single parameters. It usually stops at a suboptimal point within this valley after roughly optimizing scalp and skull values (not shown). Analysis on the final points revealed that the gradient heavily decreased until almost no error change appeared with an oscillation around a set of parameters. It had not converged and the gradient had drastically decreased.

In *Figure 5.2*, we can see the outcome of modified Newton optimization on the conductivity. Newton, in contrast to gradient descent, quickly finds the optimal 'true' values after around 20 iterations for plausible random starting values. For a low noise level (left, SNR=40dB), the optimal 'true' parameters are quickly found, while adding

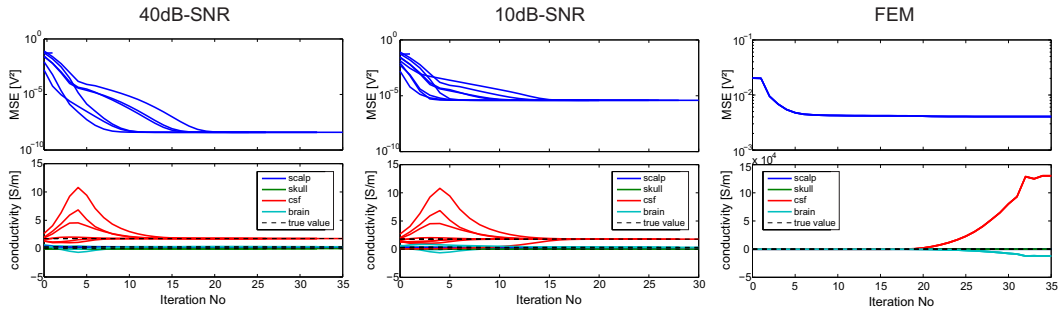

Figure 5.2: The evolution of the error for optimizations with different noise levels and reference models for an exemplary subject: while with an equal geometry head model (BEM) the optimum can quickly be found even with relatively high level of noise (SNR=10dB), for the FEM an optimum is also found quickly after which the inner parameters (CSF, brain) diverge.

stronger noise (middle,10dB) does not drastically increase convergence speed but results in a higher final error level.

In general, optimizing the conductivities of the BEM to fit the FEM simulated measurements reveals a similar behavior for the Newton algorithm as in the case of optimizing it for the BEM with noise. The algorithm reproducibly converges to a distinct set of conductivities after 12-30 iterations, while these values differ from the ones used in FEM and BEM modeling. In particular the CSF conductivity is often estimated at very low values ($\sigma \leq 0.5 S/m$) instead of the relatively high conductivity used for simulation ($\sigma = 1.65 S/m$). After this initial optimization, the values of CSF and brain conductivity drift away into highly implausible values, which are often even negative for the brain. The error still decreases slowly and no noise was added, leading to the conclusion that the algorithm counterbalances systematic geometric differences in the two models. The median source localization error using the MUSIC algorithm [Mosher et al., 1992] increases from 6.8 mm for identical conductivity values in the BEM as in the FEM model to 10.6 mm for the bEIT-optimized conductivities in 4 subjects underlining the highly implausible conductivity values.

Many multi-parameter optimizations of conductivity with lower order than Newton fail due to the shape of the error function. In particular gradient descent gets stuck in a local error valley for the values of the inner structures (CSF, brain), where the gradients decrease strongly after scalp and skull are optimized and so the optimizations are oscillating. The interdependencies of the parameters and their similar effect on the error leads to similarly low errors for very different sub-optimal ratios. Newton in contrast, modeling the quadratic part of the dependencies explicitly in the Hessian matrices, takes a higher computational load per iteration but finds the correct values with fewer steps for a corresponding headmodel. For an external validation with a more

complex FEM head model, however, optimal values are also quickly found but they over-fit the model errors/differences found from outside the head to reduce the EIT error, while the source localization accuracy decreases. Because no additional noise was added, this decrease in performance suggests an overfitting to the systematic errors in the geometry. This is most probably related to systematic differences in the model, in particular the geometry of skull and CSF due to the lower anatomical detail and smoothing of surfaces in the BEM.

5.2 Automatic Electrode Position Estimation

The position of electrodes in electrical imaging and stimulation of the human brain is an important variable with vast influences on the precision in modeling approaches and thus source localization [Wang and Gotman, 2001; Van Hoey et al., 2000; Acar and Makeig, 2013; Ollikainen et al., 2000; Michel et al., 2004].

Nevertheless, the exact position is obscured by many factors. 3-D Digitization devices can measure the distribution over the scalp surface but remain uncomfortable in application and often imprecise[Russell et al., 2005; Darvas et al., 2006]. We will now look at a new approach that uses solely the impedance information between the electrodes to determine the geometric position.

In electrical imaging and stimulation of the human brain, the interface between neural activity and the electrical hardware is mainly determined by the volume conduction properties of the individual head, the positions and size of the electrodes and their (contact) impedances. Different models with varying validity exist while the modeling error is strongly influenced by the acquisition of these variables.

Electrode positions have a huge influence on the modeling. While standardized electrode caps in different sizes are the most common approach to offer a basic positioning and reproducibility, they leave space for variations of up to centimeters. Therefor, for higher precision, electrode positions are often measured. This involves the application of 3-D digitization devices for every single subject. While this is very time consuming, the precision of the resulting data is often not satisfactory, because of (1) the imprecision of most of these devices and (2) the only vaguely known contact position and area of the electrodes.

Concerning the device imprecision (1), this is mostly better for the electromagnetic techniques such as the commonly used Polhemus [Darvas et al., 2006], while ultrasound devices offer less reliable data. Nevertheless, for most approaches the application remains unattractive due to the cumbersome acquisition process in combination with the relatively low precision.

Procedures based on moving or fixed cameras have been developed. This leads to good accuracies and a more pleasant acquisition process. However, they are based on a manual acquisition procedure [Baysal and Şengül, 2010] or automatic procedure involving a large

fixed structure around the subject [Russell et al., 2005].

The electrode position and area (2) in most approaches is only usually estimated as a point while the exact contact surface remains unclear. In gel based electrode systems, the spatial spread of this saline contact matter is smeared over the scalp, but also in water based and dry electrode systems, the exact contact position and area is not precisely known.

To ensure a good contact for most applications, impedances between electrodes are routinely measured and the contact is improved until they meet the individual needs. This impedance data is mostly obtained only between the electrode and the head which offers an estimate of the contact impedance. With small modifications to the acquisition process much more information can be gathered: by measuring the influence of injected currents on the scalp potential, the electrical properties of the whole head can be surveyed with EIT [Holder, 2005].

5.2.1 Impedance: A Measure of Electrical Distance

We can use the impedance between electrodes as an electrical distance measure and develop an algorithm to employ only this information for an automatic electrode position estimation. The measurements process can be conveniently implemented into common EEG and tDCS hardware and offers an easy-to-use solution for the automatic acquisition of electrode coordinates.

The algorithm consists of 3 main steps: we calculate the pairwise impedances between all electrodes as an electrical distance matrix (1) and use this information involving multidimensional scaling to create a Euclidean metric space (2). We then align this space to a reference and estimate the geometrical position out of a weighted interpolation from the k-Nearest-Neighbors of the electrical position (3). The weights are based on the electrical distance of the electrode to a set of base points in the reference space.

5.2.1.1 Head Model

While any common modeling approach can be used for the volume conduction properties of the individual head, a realistic 4-shell Boundary Element Method (BEM) based on individual MRIs will be used in this section. This is due to the high precision and anatomical detail in combination with a reasonable complexity. The openMEEG toolbox [Kybic et al., 2005; Gramfort et al., 2010] for BEM modeling in combination with fieldtrip [Oostenveld et al., 2011] for the interfacing is a convenient approach. The segmentation can be done involving SPMs' new_segment algorithm in combination with a post-processing. A segmentation procedure can be found in [Huang et al., 2013]. We will use 6 subjects from the OASIS database [Marcus et al., 2007] for the head model creation. The standard positions of a 64-channel Brainproducts EasyCap are aligned

and projected onto the segmentation template. It can then be morphed to the individual head involving the same non-linear warp used for segmentation. Scalp points for the reference positions are extracted from the scalp surface boundary in the relevant positions by excluding all points below a tilted coronal-transversal plane from the eyebrows to the neck. The position of the plane can be fitted manually for every single subject. Conductivities are fixed to $\sigma_{scalp} = 0.25\frac{S}{m}, \sigma_{skull} = 0.004125\frac{S}{m}, \sigma_{CSF} = 1.78\frac{S}{m}$ and $\sigma_{brain} = 0.33\frac{S}{m}$ for all subjects in this chapter.

5.2.1.2 Electrical Distance Estimation

Electrical impedances Z in a linear quasi-electrostatic situation with finite conductivities are directly interpretable as a distance metric between points, because:

- $Z \geq 0$ (they are non-negative)
- $Z(P, Q) = 0$ if and only if $P = Q$ (an impedance between 2 points is only 0 if the points coincide)
- $Z(P, Q) = Z(Q, P)$ (symmetry)
- $Z(P, R) \leq Z(P, Q) + Z(Q, R)$ (subadditivity / triangle inequality)

In the abstract case of an infinite homogeneous medium, the impedance between two points is directly proportional to the 3 dimensional Euclidean distance. Another situation where this proportionality assumption holds, is a plane, infinite conductor sheet of constant thickness with points only on one surface. In the human head, the impedances between the electrodes are dominated by the better conducting scalp which is similar to the second case. This leads to our expectations that a 3 dimensional Euclidean approach makes sense, which is primarily the case in superior scalp regions and in a local context.

In our case, we receive an EIT gain matrix $\mathbf{K}$ from the BEM head model. It maps the current injection for every electrode to scalp voltages on all electrodes: $v = Ki$. From this we can calculate the impedance Z_{ij} between two electrodes j and k by

$$Z_{(jk)} = K_{(jj)} + K_{(kk)} - 2K_{(jk)}$$

.

A measurement of real impedances inherits the difficulty of unknown electrode contact impedances. The effects can be minimized by measuring on all but the injecting electrodes [Holder, 2005]. From these measurements the pairwise impedances can be estimated. This will not be discussed in detail here.

5.2.1.3 Multidimensional Scaling

Multidimensional Scaling (MDS) is a method that first emerged in psychometrics to overcome deficits of measuring in an absolute scale [Torgerson, 1952; Shepard, 1962; Kruskal, 1964]. It takes pairwise distance measurements of a set of points in an unknown space as input and estimates the embedding of points in a multidimensional space minimizing the error in these distances. We use a metric scaling with a metric stress S as a goodness-of-fit criterion because we expect our space to be approximately metric:

$$S = \sqrt{\frac{\sum (Z_{(ij)} - \widehat{Z}_{(ij)})^2}{\sum Z_{(ij)}^2}}, \tag{5.6}$$

where $Z_{(ij)}$ is the 'real' impedance between point i and j, while $\widehat{Z}_{(ij)}$ is the resulting impedance of the model, which is an Euclidean norm of the distance in the multidimensional embedding.

The dimensionality of the resulting space is usually prefixed or chosen to fulfill some stress value criterion. However, a low number of dimensions is necessary in order to facilitate comparison of the results to the set of reference points and the inclusion of more dimensions increases the ambiguity. We will chose to generally embed into a space of dimensionality 3 (see *Section 5.2.2*).

Still, the resulting spaces of 3-dimensional scaling extracted from the different data are not directly comparable. The orientation of the coordinate system as well as the origin are ambiguous and the algorithm chooses them somewhat randomly.

We will test two different possibilities for the position of the origin: (1) the mean *centroid* of all points in the corresponding space and (2) the *electrode 'Cz'* of the head as a directly matched with the reference space.

For the definition of the different axes we will chose the following conventions:

- The first dimension points from left to right by using the direction of C1 towards C2. This direction is later adjusted to be orthogonal to the second dimension.
- The second dimension points from anterior to posterior by using the direction Pz to Fz.
- The third dimension is simply the right-hand orthogonal to the first two dimensions.

The correspondence of the output is achieved by a coordinate transform into a space fulfilling the aforementioned conventions.

5.2.1.4 k-Nearest-Neighbor Interpolation

For the transformation back to the real head geometry, we introduce a simple linear interpolation of position based on k-Nearest-Neighbors in the 3 dimensional impedance

Table 5.1: Average electrode position error over subjects with centroid and Cz as origins

origin	Average	Std
centroid	1.67cm	0.17cm
Cz	2.46cm	0.52cm

space. The nearest neighbors are found based on the euclidean distance of the electrode in the impedance space to a set of reference points on the scalp. These points are based on the vertexes of the scalp triangulation used for the head modeling (see *Section 3.2.1*). The k neighbors of each electrode with the lowest distance values are then used to calculate a weighted average $\vec{x}_{elec}$ of their geometrical position $\vec{x}_i$ in the real head space *Equation* (5.7).

$$\vec{x}_{elec} = \sum_{i=1}^{k} w_i \vec{x}_i \tag{5.7}$$

The weight is calculated from the impedance space by a linear weighting scheme *Equation* (5.8). This weighting scheme is appropriate as it keeps an influence of all neighbors while the strongest for the closest. The influence of all neighbors is appreciated because it decreases errors resulting from misalignment and distortion of the electrodes' impedance space to the reference points' space.

$$w_i = \frac{1 - \frac{d_i}{\sum_{j=1}^{k} d_j}}{k - 1} \tag{5.8}$$

Here, w_i is the influence of the position of electrode i and d_i is the distance of it to the estimated position in impedance space.

5.2.2 Estimating Electrode Locations through Impedances

The average error over all subjects is 1.67cm for the *centroid*-referenced approach, while centering on a corresponding *electrode Cz* location yields an average position error of 2.46cm. The centroid approach is significantly ($p \leq 0.01$) better than the Cz reference approach. The centroid approach is dependent on the correspondence of the scalp area included in the reference points and the electrode locations.

The results are best for superior electrode positions, which supports the hypothesis that the 3 dimensional Euclidean assumption fits best for superior areas. The largest error (12cm) is for electrode Tp7 of one of the subjects, which is one of the outer lower electrodes. It was incorrectly assigned to the other side of the head which only happened once to this extent.

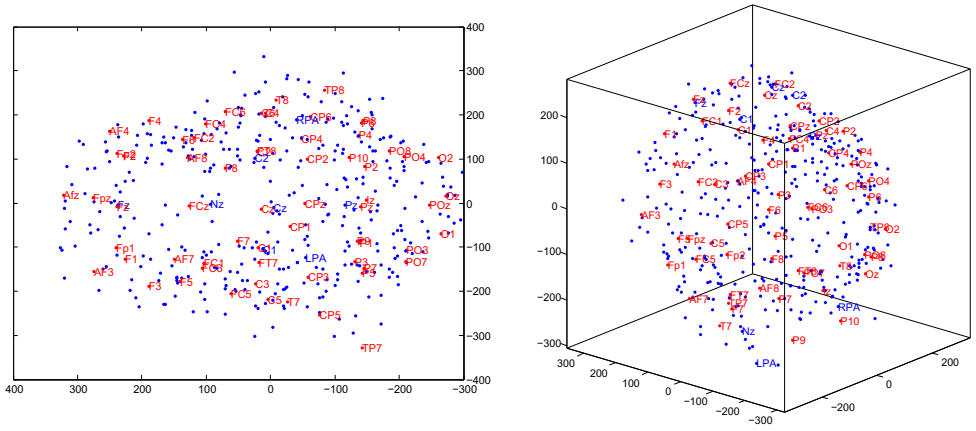

Figure 5.3: The reference locations (blue) together with the aligned electrode locations (red) in the impedance space. **left** from top and **right** in 3 dimensions.

A dimension of 3 for the impedance space seems appropriate, higher or lower numbers of dimensions increase the error for all subjects. A squared stress function improves the results for higher dimensionality but nevertheless 3 dimensions lead to the lowest errors. For a lower dimensionality than 3 the locations are not sufficiently separable with these methods, which is to be expected by the 3-dimensional geometry.

Concerning the nearest neighbor interpolation, a linear combination of the coordinate vectors of a minimum of 3 nearest neighbors improves the result significantly. A higher number has no significant effect on the average error per subject.

The results show a basic feasibility of electrode position estimation via impedance measurements. Almost no geometric or topological information of the individual head was used to estimate the geometric position of the electrodes but those of a small set of reference points. The only geometric information used is the topological location of some electrodes for impedance space alignment. The average error was 1.67cm within 6 subjects. An embedding in a dimensionality of 3 seems reasonable for the approach, while a linear interpolation from the positions of the 3 nearest neighbors is of advantage.

To fulfill the needs for precision in head modeling concerning EEG source localization or tDCS targeting, improvements have to be made. The created impedance spaces are in general of practical use but distortions are inherent. This could be avoided by introducing shape constraints in the impedance space or by introducing another method than multidimensional scaling for mapping the distances into the impedance space. Also, a better alignment procedures might improve the results especially in a higher dimensionality than 3. Here, a correct alignment should help decrease the error.

5.3 Dimensionality Reduction for Geometric Head Models

In the estimation of individual head geometries for source localization and electrical stimulation in neuroelectric investigations and applications, mostly complex geometrical models are directly extracted from anatomical images. We will now look at a novel method that uses a dimensionality reduction from thousands down to the range of tens of parameters to successfully represent an individual 4-shell BEM head model, which can successively be fitted to any kind of data from an individual head (e.g. headshape, impedances) and then used for individual head model creation. The results show that around 5 -10 components can lead to satisfactory results.

Fitting the outer head shape of an MRI to the individual, is commonly used to approximate individual head models [Darvas et al., 2006; Acar and Makeig, 2010]. [Hansen et al., 2016] have successfully used PCA as a dimensionality reduction on leadfields based on 16 different head models with a variety of conductivity values to include the estimation of the individual conductivities into their model.

In the following, we will look at a PCA based low-dimensional 4-shell representation of individual head geometries that can be easily adopted to all kinds of individual parameters, like headshape or impedance measurements, and can then be used as a basis for the electrical modeling of the individual head. The low-dimensional representation is of advantage, as it reduces complexity for optimization procedures that adopt the head model to individual parameters such as headshape warps and impedance based model generation. Some of the ideas for this approach are conceptually inspired by morphable 3D shape models of faces [Blanz and Vetter, 1999].

5.3.1 Geometric PCA for Dimensionality Reduction

For the representation of the individual geometry of 142 heads from the OASIS database [Marcus et al., 2007], we will 4-shell descriptions of the tissue boundaries for brain, cerebrospinal fluid (CSF), the skull and the scalp extracted from individual structural Magnetic Resonance Images (MRIs). These descriptions are based on single points (vertices) of the boundaries between the different tissue types and their triangular connections, which are termed meshes. From these meshes, an average head is built to transform the meshes point-wise into projection coefficients from the average head. This is done in order to linearize the description for subsequent principal component analysis (PCA) [Pearson, 1901] based extraction of the main shape manipulations as a low-dimensional description.

In a number of simulations, we will examine in a leave-one-head-out cross-validation (CV) how well different sub-approaches perform on an unseen sample and how many principle components are necessary to receive a sufficiently precise description of the

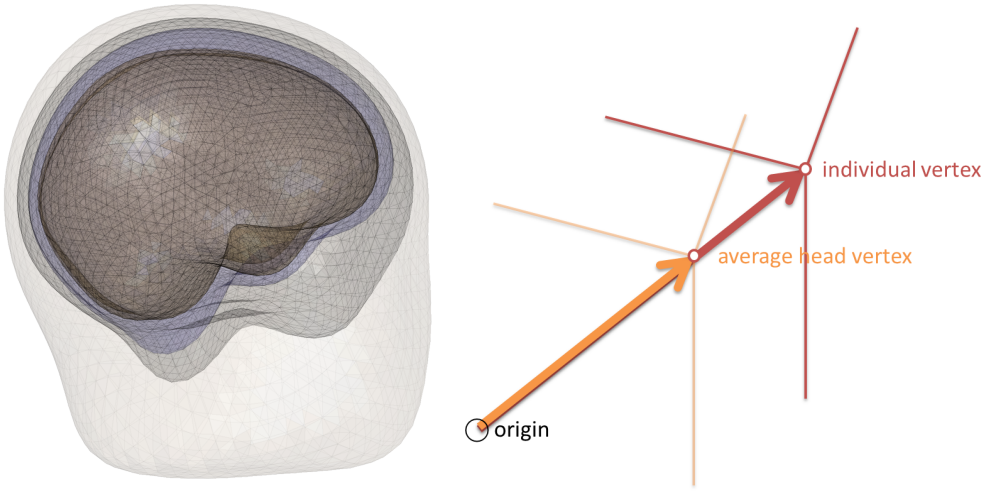

Figure 5.4: The average head from the database of 299 subjects (left) and the vector projection used to describe single subjects' vertices as a parametrization of this average head (right)

individual head geometry.

Segmentation and Meshing Segmentation of the MRIs of individual subjects are based on the *new_segment* algorithm of the SPM8 toolbox [Ashburner and Friston, 2005] and make use of an extended tissue probability map template [Huang et al., 2013]. To correct morphological errors, the Matlab function *imdilate* can subsequently be used to broaden the CSF and skull shells to a minimum thickness of 3 voxels, and the scalp shell to a minimum thickness of 4 voxels. Surface meshes are extracted using the ft_project_mesh function of the fieldtrip toolbox [Oostenveld et al., 2011]. Used here are 1922 nodes for every surfaces (scalp, skull, CSF, and brain). Nodes located more than 25mm away from any neighbor were relocated to the geometric mean of all neighbors. The meshes were smoothed using vertex-based anisotropic flow [Zhang and Haniza, 2006].

Data A co-registered average across 3 to 4 individual MRI scans with 1mm resolution of 299 non-demented subjects from the OASIS database [Marcus et al., 2007] has the advantage of a high image quality and will thus be used. The subjects' ages this chapter range from 18 to 94 with an average of 45 years consisting of 113 men and 186 women.

Cross-validation To avoid over-fitting and evaluate generalization performance of the procedure, we will conduct a leave-one-out cross-validation over subjects. Hereby, we

first extract average head and principal components from 317 subjects and then used the linear least squares reconstruction (LLS) on the left out.

5.3.1.1 Average Head Extraction and Projective Representation

In order to obtain an easy and close-to-linear representation of every single head of the database, an average head is produced and every single head is represented as a linear projection of the points of this average head. The extraction of the tissue interfaces/surfaces as triangular meshes is advantageous for this subsequent linear vector projection because the points lied more or less on single lines from the origin for analogously extracted meshes. The average head is built by averaging all single points over all subjects. The resulting average head can be seen in *Figure 5.4* together with a visualization of the vector projection used.

5.3.1.2 PCA Extraction & LLS Reconstruction

Principal Component Analysis (PCA) was designed for identifying linearly uncorrelated principal directions in a set of correlated variables and can thus be used as a dimensionality reduction. The centering around the average head and the projective description ensures the variables for PCA to be average-free and to consist of linearly correlated variables, both prerequisites for it to work well. In order to investigate the necessary complexity of the model but also for possible applications, different approaches for extraction and reconstruction are performed. For individual head geometries, the different PCA based representations were then fitted individually by a linear least squares (LLS) fit of the euclidean-distance between model and reconstruction involving different numbers of PCA components.

5.3.2 Lower Dimensional Head Approximations

5.3.2.1 Shell-Wise Approach

In the shell-wise extraction of the individual geometries, PCA is applied to every single tissue surface (scalp, skull, CSF, brain) individually, leading to 4 individual dimensionality reductions of 1922 variables. Also the LLS reconstruction is done individually for every single shell.

The shell-wise extraction as the most detailed leads to the lowest errors compared to the other approaches, see *Figure 5.5*. It represents every tissue surface by a separate principle component decomposition. Already with 2 principal components (PCs) per surface (leading to 8 PCs for the 4 surfaces in total), the average geometric errors per head are mostly below centimeters and above 10 PCs used for reconstruction the approximation increasingly approaches mm-precision (50 PCs).

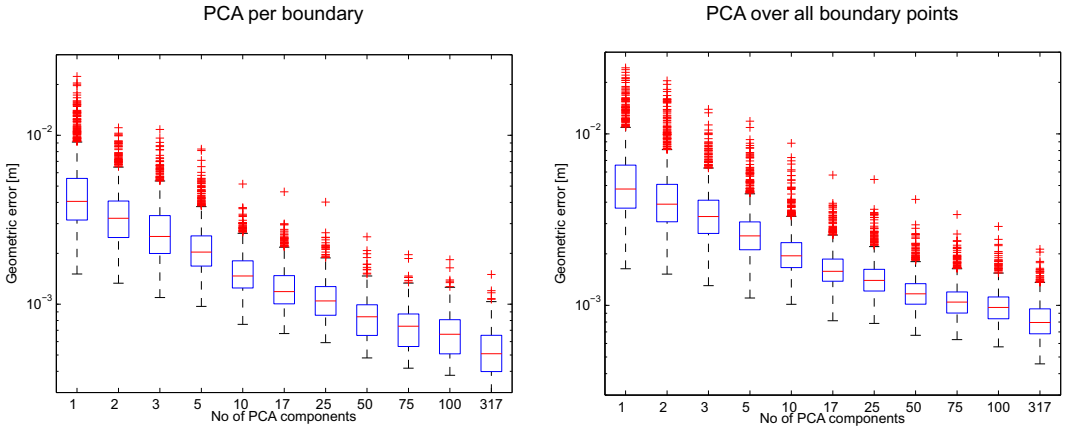

Figure 5.5: The average per subject geometric error of the PCA based approximation to the real head geometry depending on the number of components used. On the left, one PCA per boundary is used so the actual number of components is four times higher. A clear decline can be seen reaching a sub-millimeter for most boundary points at around 50 components per surface (200 in total). On the right, PCA is performed over all boundaries at once. Also, a clear decline can be seen but errors are in general higher than for the shell-wise approximation.

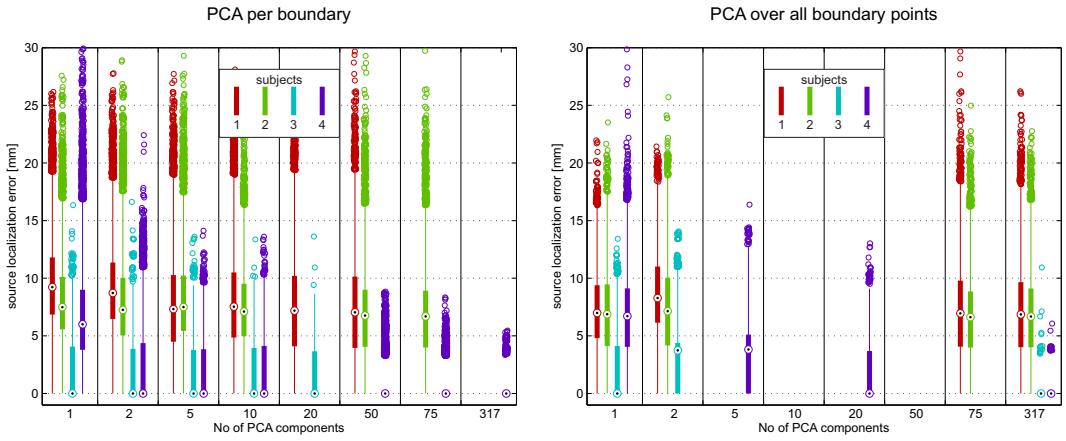

Figure 5.6: The per subject source localization error in mm of the PCA based approximation to the real head geometry depending on the number of components used. On the left, one PCA per boundary is used so the actual number of components is four times higher. While for the one PCA per boundary, many head model creations failed due to intersecting meshes in the BEM, low errors can be reached quickly. For the PCA over boundaries on the right, the models are more stable and errors are lowest already with 5 components.

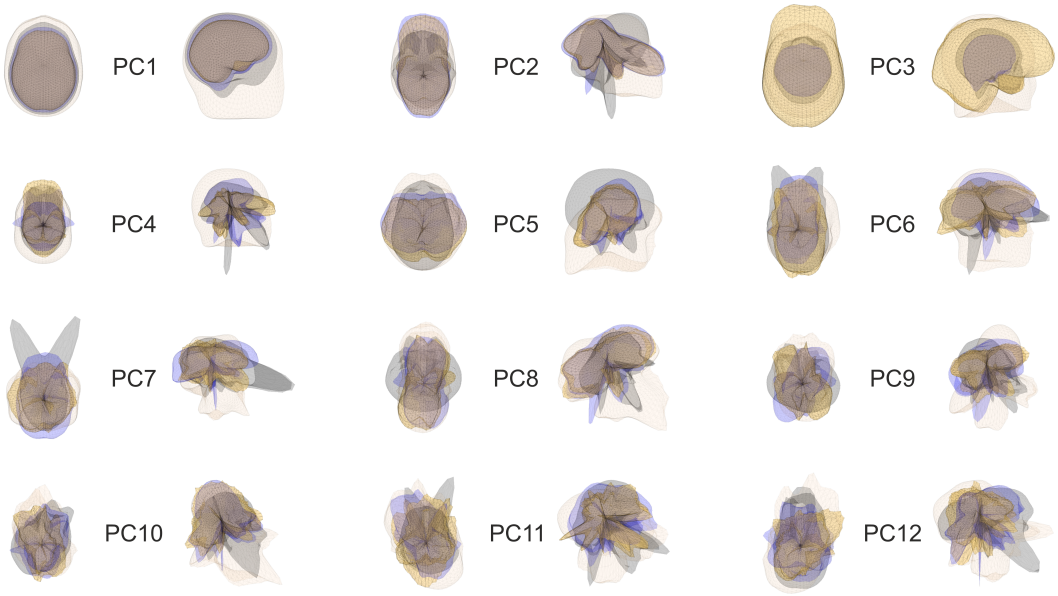

Figure 5.7: The first 12 PCA components sorted by explained variance each from two different geometric directions for the shell-wise extraction: one PCA per boundary was applied but they are plotted as an overlay here.

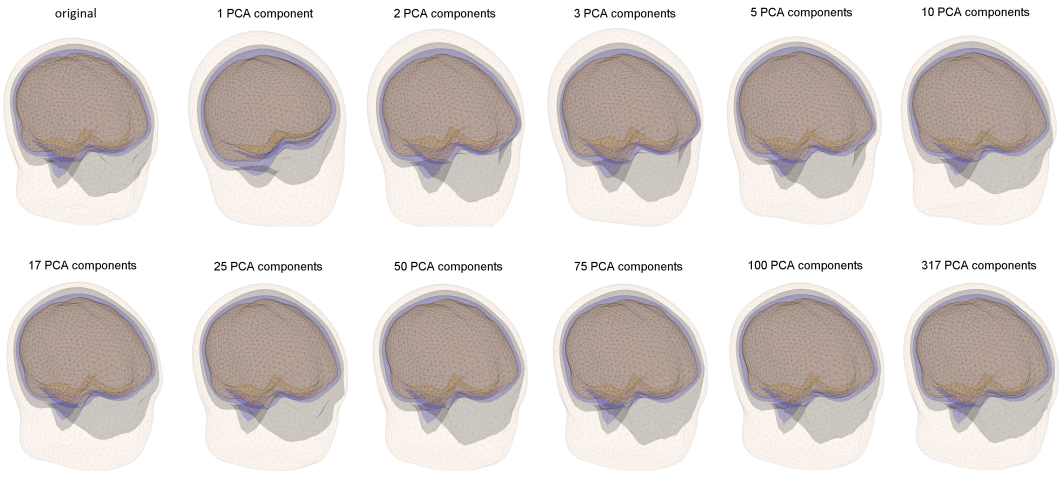

Figure 5.8: The reconstruction of an exemplary subject (OAS1_0408_MR1) for the shell-wise extraction: one PCA per boundary was applied.

If we look at the first 12 PCs per surface (48 in total) explaining most geometric variance in *Figure 5.7*, we can see that the different surfaces can strongly vary in shape and size. From the 5^{th} PC on the shapes are getting more abstract and less comparable to actual head geometries.

For single subjects (here OASIS1_0153), we can find in *Figure 5.8* that from 2 PCs per boundary (8 in total) on, a meaningful head geometry was obtained. Using more components constantly brings the approximation closer to the original head and from around 25 components per boundary most important details like local tissue thicknesses and overall shape appear well approximated.

5.3.2.2 Multi-shell Approach

In the multi-shell extraction of the individual geometries, PCA is applied to all tissue surfaces at once (scalp, skull, CSF, brain), leading to one single dimensionality reduction on the $4 \times 1922 = 7688$ variables. The LLS reconstruction is done for all shells at once.

The multi-shell PCA approach leads to a very similar picture in *Figure 5.5*, although, error levels appear slightly higher in general.

Investigating the components (not shown), we can see that already the first two PCs look like real head geometries and overall the PCs look more concentric concerning the layers than with the single-layer approach. Also, PCs are more abstract from the $8th$ component on, later than with the single-layer PCA approach. The first component is very similar to the average head. From around 5-10 PCs on, the most important details like skull and CSF seem well approximated.

5.3.2.3 Headshape Reconstruction

In this approach, PCA was applied analogously to the multi-shell extraction, but the LLS reconstruction is done based on the outermost shell only - representing the head shape. This is done to have an application scenario similar to common headshape-based warps [Darvas et al., 2006; Acar and Makeig, 2010]. The shape based reconstruction is as such only applicable to the multi-shell approach as we need a linked set of components for all surfaces.

For basic evaluation of a possible application of the low-dimensional approach, we can examine in *Figure 5.9* how well the heads are approximated if we only adjusted the headshape. The first two components lead to a very similar error distribution as the multi-shell approach but with more components the decrease in error is lower.

5.3.3 Remarks on Lower Dimensional Head Models

The approach to dimensionality reduction in individual head geometry modeling is successful, leading to low source reconstruction errors for a representation with around 5-10

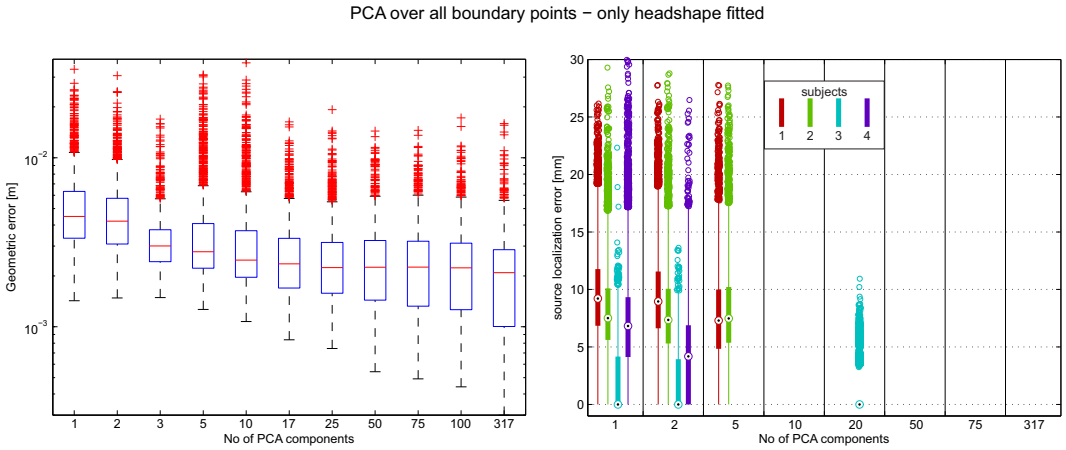

Figure 5.9: The average per subject geometric error and the per subject source localization error of the PCA based approximation using the headshape only: the PCA extraction is based on all boundaries at once but the approximation is done based on the most outer boundary only - depicting the headshape. Also here, an improvement in geometric error can be found but sub-millimeter ranges are not reached. The source reconstruction error is with very low number of components relatively higher compared to the other approaches, but also quickly decreases. Reconstructions are less stable as with more components, meshes mostly intersect.

parameters. Compared to the $4 * 1922 = 7688$ boundary points this is a very low number of parameters and thus a strong dimensionality reduction. The median errors remain below centimeters for all approaches and already with one component. This is an effect of the anatomical similarity of human heads in general. While applying one PCA per boundary has the potential to represent more anatomical detail and thus lower errors, the PCA over all boundaries might avoid other errors like interfering surfaces and is likely to be more suitable for applications in head geometry optimization. Also, the source reconstruction error for the four subjects is lower after fewer iterations. Fitting only the headshape works but has the inherent problem of not knowing the inner tissue distributions and thus leads to higher geometric errors for the inner structure (skull, CSF, brain), but the source reconstruction is less affected and leads to acceptable values.

5.4 Linear Leadfields

In order to overcome the weaknesses of generic head models, mostly individual magnetic resonance images (MRIs) are taken, segmented and a head model is built based on

the individual anatomy. This is very intense in cost and labor. In this section, we will investigate existing approaches including a leadfield average and, additionally, a new method that involves only the information of impedance measurements from the individual subject to create an individual headmodel as a linear approximation from a database.

The linear approximation approach uses a database of leadfields and EIT gain matrices to approximate an individual head model based on electrical impedance measurements only. We have introduced the different methods used for anatomical extractions from MRI and head model generation in chapter *Chapter 3*. We will use the extraction based on [Huang et al., 2013] described in *Section 3.2.1.1* and the models in *Section 3.2.1.3*.

The findings of [Valdés-Hernández et al., 2009] that the average leadfield works comparably to a non-linearly warped MRI in combination with the success of EIT based approaches for conductivity estimation has lead to the idea of adopting an average leadfield through EIT. The novel linear approach refines an average head model by an optimal linear combination of precalculated spatially aligned leadfields - weighted by the optimal weights over the subjects from a linear least-squares optimization for EIT measurements. It stems from the group of approximate head models but involves impedance measurements similar to EIT for the head model individualization.

This approach is different to common EIT approaches that mostly directly solve the inverse problem by Thikonov regularized least-squares approaches for the individual compartment (mostly voxel-wise) conductivities [Holder, 2005].

We will also compare the approximation to other common methods, namely the New York Head ICBMny [Huang et al., 2015], a non-linearly warped template BEM [Tadel et al., 2011; Acar and Makeig, 2010, 2013], constant skull thickness 3-shell BEM [Oostenveld et al., 2011; Tadel et al., 2011], realistic 3-shell BEM [Akalin-Acar and Gençer, 2004] and realistic 4-shell BEM [Acar and Makeig, 2013].

We will first inspect the main ideas of the linear approximations through impedance measurements which are characterized in *Section 5.4.1*, then describe the anatomical data used for the approximations in *Section 5.4.2.1* and four separate MRIs used for evaluation in *Section 5.4.2.2*. After this, the source and EIT simulations used for validation will be introduced in *Section D.3*.

5.4.1 Theory on Linear Approximations of Leadfields

Up to now we have only described the EEG and the EIT approaches in parallel and their correspondence. We will bring them together to follow the idea of avoiding costly MRI and build an individual head model on the basis of impedance measurements only. The relationship between EIT gain matrix and EEG leadfield matrix is highly non-dimensional and complex. Although they both form linear relationships between currents and the sensor voltage, their currents originate in different structures: one from within the brain (EEG) and one from scalp current injections (EIT). Additionally, the

dependence on the underlying hidden geometric parameters is strongly non-linear. E.g., to estimate a leadfield matrix L directly from the corresponding gain matrix K, we could solve the corresponding inverse EIT problem first to estimate the geometry and conductivity for the individual subject after which we could solve a BEM or FEM model for the EEG sources to obtain the leadfield matrix **L**. This solution is, however, tedious and error-prone.

The presented approach to finding the leadfield corresponding best to our EIT measurements is based on a simpler, linear approximation. The fact that the average leadfield of a population of 142 subjects was found to provide a good approximation [Valdés-Hernández et al., 2009] for individual heads however leads to the expectation that a weighted average may even improve upon that.

The volume conduction modeling involving the Boundary Element Method (BEM) was described in *Section 3.2.1.3*. Note, however, that the procedure can be used with any common head modeling approach representable in this linear form as it acts solely on the gain matrices.

The EIT gain matrix **K** (*Equation* (3.6)) and the leadfield (or EEG gain) matrix **L** (*Equation* (3.7)) were introduced in *Section 3.2.1.5*, which have been built from the BEM head models.

By modeling the head for every single subject in our database, we get a database of matrices as solutions for EEG and EIT case:

$$\mathcal{K} = \{\mathbf{K}_i\}$$
$$\mathcal{L} = \{\mathbf{L}_i\}$$

$\mathbf{K}_i$ has dimensions NoElec×NoElec and $\mathbf{L}_i$ NoElec×NoDipoles×NoGeometricalDimensions.

First, we need to solve a (regularized) linear least squares problem to approximate the measured EIT voltages (with known currents) by a linear combination of gain matrices $\mathcal{K}$ from the database.

$$\hat{\mathbf{w}} = \arg\min_{\mathbf{w}} \sum_{k=1}^{N} \left(V_k - \left(\sum_{i=1}^{M} w_i \mathbf{K_i} \right) \cdot \mathbf{i_k} \right)^2 \tag{5.9}$$

Here, $\mathbf{w}$ is the vector spanned by the space of all individual weights w_i and $\hat{\mathbf{w}}$ denotes the optimal vector. N is the number of measurements and M is the number of models in the database. The individual injection current pattern of measurement k is given by the current pattern vector $\mathbf{i}_k$ and the voltage elicited by this injection is measured as a scalp pattern V_k. In general, we can measure all of the entries of **K** by injecting currents through different electrodes. If we inject through a pair (one source and one sink) of electrodes at a time only, we can successively measure the scalp voltage on all sensors V_k for all or a subset of possible pairs $\{k\}$ of injection sites. But also more complex injection

patterns are possible, e.g. to reduce measurement time or to improve the measurements [Holder, 2005].

To approximate the matrix $\mathbf{K}$ from the measurements, we seek the weight vector $\hat{\mathbf{w}}$ that is the minimizer of the linear least squares problem in *Equation* (5.9). This has a unique and analytical solution.

The approximation $\hat{K}$ is then obtained as

$$\hat{\mathbf{K}} = \sum_{i=1}^{M} w_i \mathbf{K}_i$$

Having an estimate of the EIT gain matrix $\mathbf{K}$, we are still left with the task of estimating the leadfield $\mathbf{L}$ which is required for EEG source imaging and tCS targeting. In general, we have to assume a non-linear relationship between $\mathbf{K}$ and $\mathbf{L}$:

$$\hat{\mathbf{L}} = f\left(\hat{\mathbf{K}}\right)$$

In the linear leadfield approach, we now consider a linear approximation for simplification, under which $\hat{\mathbf{L}}$ is obtained as a linear combination of the lead fields $\{\mathbf{L}_i\}$ from the database, where the optimal weights are given by the $\hat{\mathbf{w}}$ previously optimized through electrical impedance tomography:

$$\hat{\mathbf{L}} = f\left(\hat{\mathbf{K}}\right) = f\left(\sum_{i=1}^{M} w_i \mathbf{K_i}\right) \approx \sum_{i=1}^{M} w_i f_i\left(\mathbf{K_i}\right) = \sum_{i=1}^{M} w_i \mathbf{L_i}.$$

Negative weights on gain matrices (lead fields) flip the polarity of the scalp potentials and therefore amount to subtracting gain matrices (and lead fields) from another in order to potentially achieve better approximations of the EIT measurement. However, as all anatomies in the database were spatially aligned to the same template, corresponding gain matrices and lead fields are expected to be roughly similar in particular in polarity. We must therefore assume that sign flips occurring due to negative optimization weights are more likely results of over-fitting than of genuine anatomical differences. As a consequence, we will constrain all weights to be non-negative. The details of the non-negative linear least squares algorithm can be found in [Lawson and Hanson, 1995]. It solves *Equation* (5.9) with the constraint that $w_i \geq 0$. A comparison of the performance of unconstrained and non-negative weights based approximations is provided in *Section 5.4.3.1.*

5.4.2 Anatomical Data Extraction

The realistic head models for validation are built with another pipeline and method than BEM in order to avoid inverse crimes using the exactly same model for validation as for modeling. Different types of head model approximations for the evaluation subjects

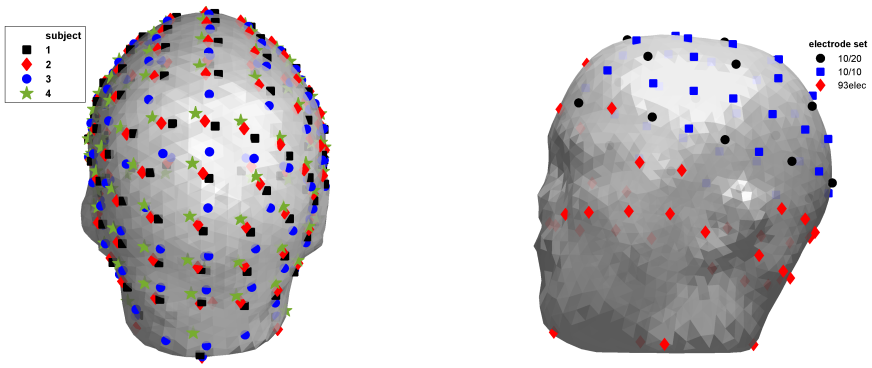

Figure 5.10: **left** The 4 different electrode sets from the external validation subjects warped onto one of the heads from the database (OAS1_0001) **right** The first electrode set on OAS1_0001 with the different electrode subsets.

were also constructed, in order to have the results directly comparable to other common approaches within the same experimental settings. The non-linear warp, estimated for segmentation between the individual and the template, is the basis of source locations for all head models, as this is guarantees brain-topological correspondence. For all of the other common approximations except for the ICBMny head (see below), also electrode locations are based on this warp. The ICBMny as a template based procedure does not know the individual MRI, similar to the average and linear leadfields. In practice, these usually rely on a label match criterion and, as a consequence for the experiments, the effect of knowing the electrodes (similar to [Valdés-Hernández et al., 2009]) and not knowing them for the average and linear approximations is investigated. As measuring the ground truth (the true fields within the head) is infeasible in healthy humans, a highly detailed FEM model is expected to be the closest possible to reality, but is expensive to calculate. Another set of 4 heads for an external validation with an FEM model and the 4-shell BEM modeling approach for the linear and average leadfield approximations is used to avoid the calculations of FEM head models for the whole database and due to the favorable format for the approximation.

Hereby, we first compare the resulting leadfields directly with different error measures, after which we compare the simulated outcome of EEG source localization for pseudo-EEG and a simulated tCS targeting application.

The details of the anatomical extractions can be found in the Appendix D of which follows an overview, here.

Source Locations Only in the investigations about the necessary size of the database and conductivity, we will use the BEM as a comparison. The location of neuronal sources, is extracted from the gray-white matter boundary as described in *Section 3.2.1.4*. Con-

tributions from all over the brain might be generally found in the EEG and, thus, we will investigate the behavior of the leadfields in the whole brain but the analysis will also be carried out on source locations limited to the cerebrum close to the cortical surface. By this, of the original 3990 locations extracted from the gray-white boundary 149 locations were excluded because they ended up out of the brain structure for some of the subjects in the database. From the resulting, 2173 source locations are used as locations for the closer analysis limited to the neocortical surface (see yellow dots in *Figure 3.5*).

Electrode Positions In EEG, electrodes are the sensors in contact with the biological tissue and their geometric position determines the amplitude of the signals from different areas of the brain. This is the basis of EEG source localization: if we record at different locations simultaneously, we can reconstruct where in the brain the signal originated. In EIT and tCS targeting the sensors are not limited to measuring voltages but are also the location of current injection. Also in these cases, the locations of the electrode are an important factor.

In order to obtain an optimal correspondence over the subjects' sensor positions, warped electrode locations from an aligned electrode set in the template space to the individual head in a similar process as the source locations are used. This warp is based on the same non-linear warp used in the source position warp that stems from the segmentation. By this, all electrode positions are in good correspondence with the whole head segmentation and the source grid in particular as the warp is by definition smooth over the whole head. Projections of source and electrode locations onto the head surface end up in a similar topology before and after the warp. The approach is similar to [Valdés-Hernández et al., 2009].

In *Figure 5.10* on the right we can observe the positions of the different electrode subsets used for evaluation, which are:

- **10/20** A set of 19 standard electrode positions, displayed in black dots.
- **10/10** A set of 57 standard electrode positions, the union of black dos and blue squares.
- **extended 10/10** A set of 93 standard electrode positions, all the colored symbols together.

The ICBMny is evaluated for corresponding and known electrode locations and the realistic setting of a simple electrode label match. For the corresponding setting, the closest electrodes in the template head model compared to the ones after the warp from individual MRI to the ICBMny template were chosen according to a geometric nearest neighbor criterion.

EIT Simulations The following procedure is done to simulate the scalp potentials produced by external current injection through the electrodes: All possible pairs of the electrode set are combined by setting a positive current of $100\mu A$ at one electrode and a negative return current of the same amount at another. Because the whole model is based on a linear current-voltage relationship (Ohm's law), the current amplitude actually introduces relative effects and the absolute value is not important for the simulations. Different noise levels are simulated including a much stronger noise level of Gaussian noise on the measurements than is to be expected leading to an SNR of 20dB (10:1) to evaluate the maximally expected influence. Expected is around 60dB (see *Section 3.1.2.4*). The simulated EIT measurements of the highly detailed FEM head model are used to construct a both linear least-squares and non-negative linear least-squares approximations, while the database is built involving BEM models.

5.4.2.1 Approximation Subjects

An open-source MRI database for the creation of a database of head models is used which is the same as in section *Section 5.3.1.* Only 142 single MRI sessions of the first 5 discs are used. There are 58 male and 84 female subjects of age range 18-90.

The size of the head model database is chosen to be big enough to be a representable collection of different head geometries and to keep memory usage low enough for computation. In order to determine the necessary amount of subjects, the amount of head models used for the linear approximation was varied and error measures observed in an internal validation within the database. These error measures saturated at the inclusion of around 100-150 subjects as you can find in the results *Section 5.4.3.4.* 41 subjects with a Clinical Dementia Rating (CDR) higher than 0 were excluded to avoid demented subjects in the data because demented subjects were found to have a alternation in gray and white matter distributions (e.g. [Brun and Englund, 1986; Schmahmann et al., 2008]). Additionally, the MRI of subject OAS1_0121 was corrupted which lead to an additional exclusion of this subject.

5.4.2.2 Validation Subjects

MRIs ($1mm^3$ isotropic resolution, T1-weighted, 3T magnetic field) of four healthy individuals, all Caucasian male, age range 27–45) were acquired. From these MRIs individual FEM models were built and then used for neuronal source and EIT simulation. The details of this modeling are described in *Section D.3.* Note that this is a different pipeline than the head modeling used for the approximations.

From the external 4 MRIs individual FEM models are built after a manually refined semi-automatic segmentation procedure. The EEG leadfields and EIT gain matrices are solved using commercially available Abaqus software with a $1mm^3$ resolution. The details of this model extraction can be found in [Huang et al., 2015]. For this model,

specific conductivities are $\sigma_1 = 0.465\frac{S}{m}$ for the scalp, $\sigma_2 = 0.01\frac{S}{m}$ for the skull, $\sigma_3 = 1.65\frac{S}{m}$ for the CSF $\sigma_4 = 0.126\frac{S}{m}$ for the gray matter and $\sigma_4 = 0.276\frac{S}{m}$ for the white matter.

5.4.3 Analysis of Results

5.4.3.1 Comparison To Other Approximations

For the validation of the method, different error measures are analyzed based on the similarity to realistic FEM head models obtained from individual MRIs. The common error measures Relative Difference Measure (RDM) [Meijs et al., 1989] and the Logarithmic Magnitude Difference Measure (lnMAG) [Güllmar et al., 2010] of the linearly approximated leadfields and other common methods are compared to a highly detailed FEM leadfield. RDM measures the relative difference and is commonly related to source localization error, while the lnMAG give information about the similarity in amplitudes two signals.

To show the relevance in application, exemplary source localization is carried out and compared over the different methods, within and between subjects. The error measures are described in detail in Appendix *Section C*.

The following different head modeling approaches will be performed:

- **ICBMny** High detail FEM model based on the ICBM152 anatomical template (a non-linear average of the MRI of 152 adult human brains) with six tissue types (scalp, skull, CSF, gray matter, white matter, air cavities) at $0.5mm^3$ resolution. [Fonov et al., 2009, 2011; Huang et al., 2015]. The electrode positions were not adjusted but only assigned through a label match criterion, as this is the common case.

- **warp4** Similar to [Tadel et al., 2011; Acar and Makeig, 2013], a non-linearly morphed MRI of the Colin27 head is produced within *brainstorm* [Tadel et al., 2011] from the positions of a standard 10/5 electrode set with 343 electrodes and then applied the 4-shell BEM modeling approach (described in *Section 3.2.1.3*) to create a head model based on this. The reason for the assumption of known electrode locations, here, is that they were measured to obtain the morphed template MRI. The specific conductivities are the same as in any of the 4-shell BEM: $\sigma_1 = 0.465\frac{S}{m}$ for the scalp, $\sigma_2 = 0.01\frac{S}{m}$ for the skull, $\sigma_3 = 1.65\frac{S}{m}$ for the CSF and $\sigma_4 = 0.33\frac{S}{m}$ for the brain.

- **const3** In this model, we use straight standard function from *fieldtrip* [Oostenveld and Praamstra, 2001] to produce a constant skull thickness 3-shell head model with a skull thickness of $4mm$. The specific conductivities are $\sigma_1 = 0.465\frac{S}{m}$ for the scalp and $\sigma_2 = 0.01\frac{S}{m}$ for the skull, $\sigma_3 = 0.33\frac{S}{m}$ for the brain.

- **real3** This head model simulates the absence of CSF in the geometry by simply taking the same boundaries for scalp and skull as the realistic 4-shell model and using the outer CSF boundary as a brain boundary. The specific conductivities are the same as for the **const3** constant skull thickness approximation.
- **real4** 4-shell BEM model involving the realistic segmentations described in the methods section, 1922 vertexes per surface triangulation.
- **average4** Average leadfield of the database.
- **linear4** Linear leadfield from the database with non-negative constrains, the EIT measurements are simulated using the FEM head model.
- **linear4nc** Linear leadfield from the database without constraints in the approximations, the EIT measurements are simulated using the FEM head model.

We will now introduce the different experiments we have investigated:

Corresponding Electrode Locations In these simulations, the error made by imprecise correspondence of electrode locations, head model and source locations was out of focus: leadfields and EIT gain matrices are constructed for every subject of the database involving the 4 different electrode sets of the validation subjects. The geometrical transformation from the template to the individual head is based on the non-linear warp used in the segmentation. This leads to corresponding segmentations, electrode and source locations for and between the heads of the database and the subject.

With exactly known electrode locations - transferred to every single head from the database - the non-negative linear approximation leadfield (linear4) performs comparable in source localization to the individual realistic 4-shell BEM (real4) over all but one subjects. The unconstrained linear approach (linear4nc) leads to the least reliable approximations. It has more errors than the warped template head model (warp4). The non-negative weight vectors has sparse solution with about 10% of non-zero entries, while the unconstrained least squares weights are almost fully non-zero (94%). The ICBMny head model has lowest amplitude error and is closest to the BEM4 in RDM but produces higher error distributions in source localization.

The ICBMny performed best concerning magnitude (lnMAG) followed by the non-negatively constrained linear lead fields. There are least bias in the amplitudes of the ICBMny leadfields (see *Figure 5.11*) and lowest overall quantiles measured from the median. The next lowest overall errors are produced by the real4, followed very close by

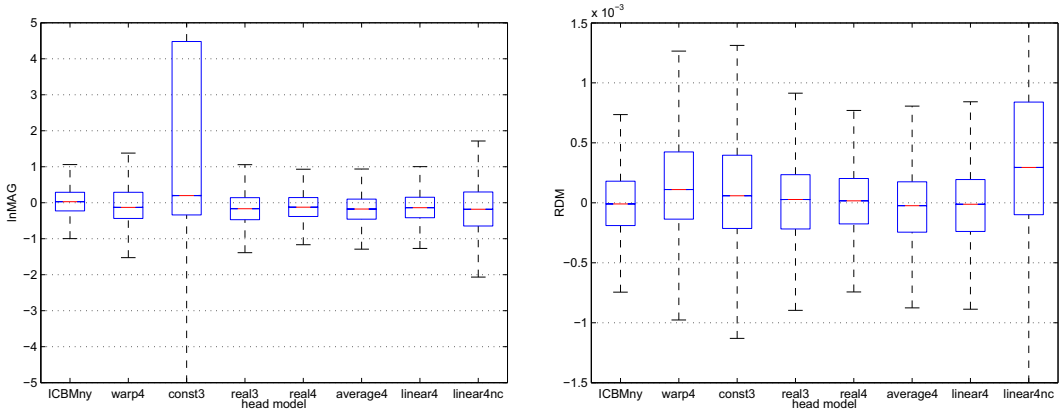

Figure 5.11: The lnMAG and RDM in different head model approximations over 4 subjects. The different abbreviations for the approximations are explained on page 105.

the linear4. The average leadfield attains mostly smaller amplitude values (lnMAG<0) then the FEM, while the values attained by the unconstrained linear approach tend to be too big.

In RDM, real4 and ICBMny lead to very similar results. The next smallest values are produced by average and linear non-constrained leadfield, while the differences between the two are not very pronounced but visible. The bias of both of the linear4 is smaller but the quartiles are very similar. This time the const3 is further off than the linear4nc followed by the warp4. Again, ICBMny is the closest to the real4 head model.

We can also see behavior very similar to the individual 4-shell BEM in the source localization for the non-negative linear approximation (figures *Figure 5.12*). The average has the next lowest overall distribution of source localization errors after which the real3 is followed by similar results of the const3 and ICBMny.

At a single subject level (*Figure 5.13*), the source localization error is very similar for the average leadfield for all the subjects, and in the non-negative linear approximation for all but one subject. The non-negative linear approximation is slightly worse than the average leadfield for subject 4, while it improves the source localization for the others. Still all the approximations are better than 3-shell individual head models across subjects (compare to *Figure 5.12*).

Examining the spatial distribution of source localization errors for the non-negative constrained linear leadfield in *Figure 5.14* of *Chapter 3*, we can see that locations close to the cortex surface had lowest error while deeper structures had bigger errors. Mostly due to the smoothed boundaries, the model errors in the 4-shell BEM we used were biggest in lower skull regions and facial areas, which also lead to biggest mismatches in source location estimation.

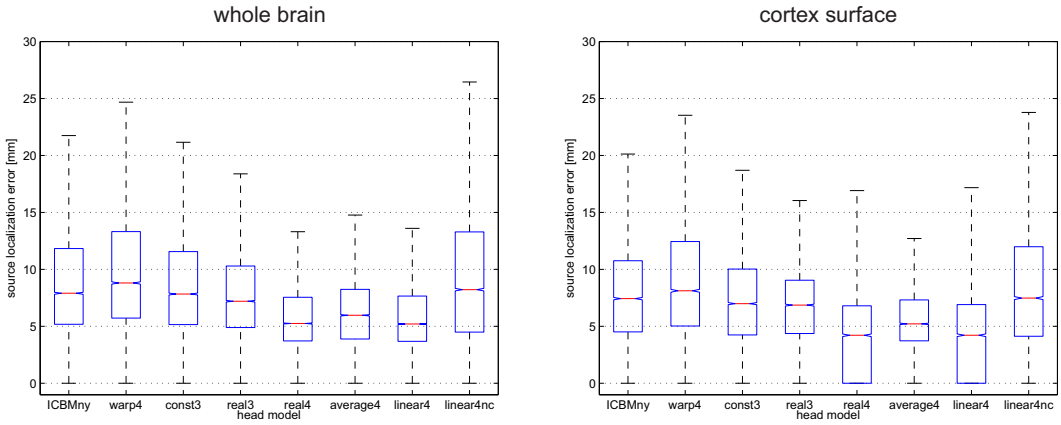

Figure 5.12: The performance of source localization in different head model approximations over 4 subjects (57 electrodes 10/10 system) for the whole brain (**left**) and the cortex surface only (**right**): we can see a decrease in source localization error for increased anatomical realism in the individual head models. Interestingly, the ICBMny is very realistic but not individualized and it still produces lower errors than warp4. Looking only at the cortex surface, the overall errors decreased. The different abbreviations for the approximations are explained in in the box on page 105.

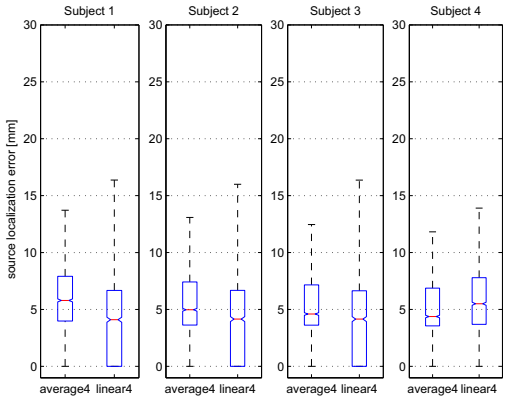

Figure 5.13: The performance of source localization in different head model approximations involving the corresponding electrode sets for the 4 subjects individually in the cortex surface (57 electrodes 10/10 system): the linear4 approach decreases most errors compared to the average except for subject 4.

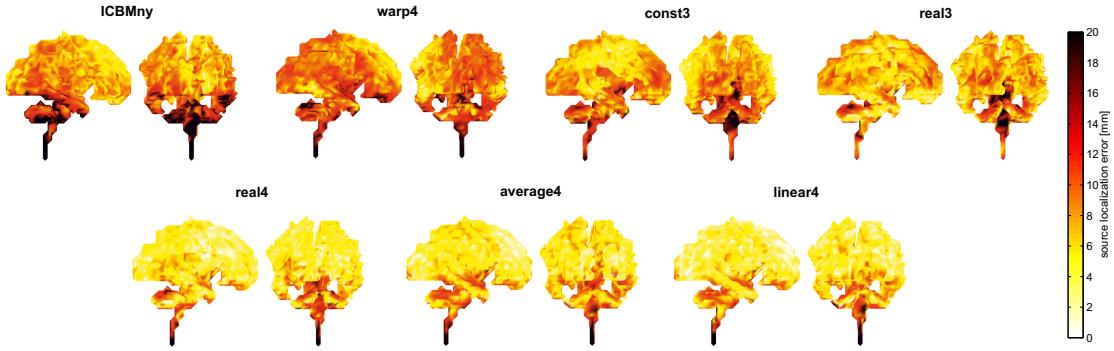

Figure 5.14: The topographies of source localization error on the white-gray matter boundary for the different approximations as an average over 4 subjects (57 electrodes 10/10 system). Note that the values are clipped above $20mm$ for visualization.

The procedure of creating corresponding source and electrode locations through a common template provides similar leadfields and EIT gain matrices in head modeling across all approximations. All median errors in source localization lie around 10mm, RDM are around 0 and stay within the same scale except for the constant skull thickness 3-shell BEM. It leads to the strongest amplitude errors. A 4-shell BEM approach performs best - compared to the other common approximations avoiding highly detailed FEM simulation - in source localization and other error measures.

Imprecise or Partial Correspondence of Electrode Locations In general, the exact electrode location for the future subjects is not known when we create the leadfields for the database. In the last section, we had assumed known and corresponding electrode locations between the subject and the head models in the database. The correspondence, which is also present between the electrode positions and the source locations, is based on the same known non-linear warp that is applied to co-register anatomies. This is done in order to measure the quality of the lead field approximation. The map however requires knowledge of the individual electrode locations and MR images, which is exactly what we seek to avoid in practice.

In a more realistic setting within this section, we will assume electrode locations on the individual head that are set under the same standard 10/5 policy for every subject. Every electrode set from one of the evaluation subjects is warped to the individual database head. This results in 4 different leadfields per database subject.

For the ICBMny, we will test a similar setting, where only the electrodes assigned by a label match between the template and the individual electrodes are used and compared to a nearest neighbor criterion.

In the following figures, we compare:

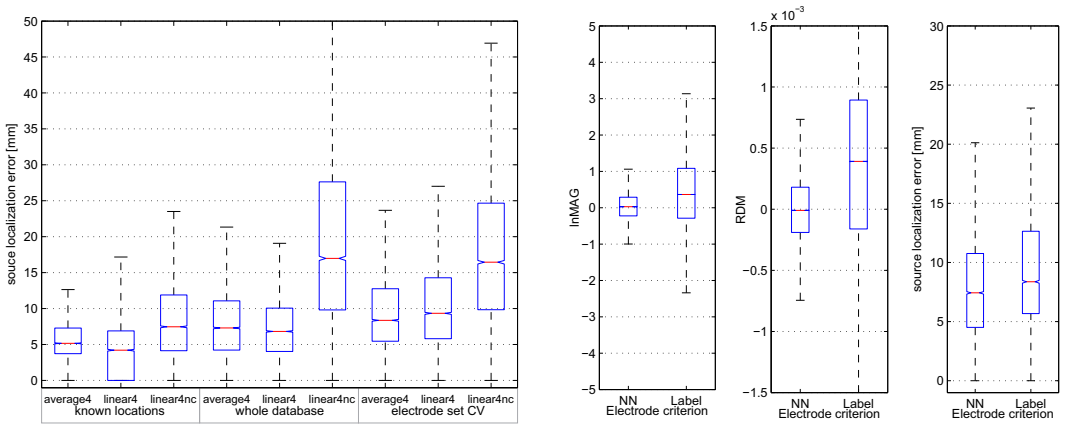

Figure 5.15: **left** The performance of cortical surface restricted source localization in different head model approximations and electrode sets over 4 subjects (57 electrodes 10/10 system). **right** The results for the ICBMny generic head model with geometric nearest neighbor (NN) and with label electrode match over 4 subjects (57 electrodes 10/10 system). The correspondence of electrodes is essential for precise source localization.

known locations - 142 heads with corresponding electrode set
whole database - All 4 electrode sets within the database (4x142=568 leadfields)
electrode set CV - 3 of 4 electrode sets: the test head excluded (3x142=426 leadfields).

Source localization was affected by the different databases. For the approximation *whole database* based on all 4 electrode sets (so including the corresponding one) average4 and linear4 were less precise but the linear approximation still improved results compared to the average. With the corresponding set excluded (*electrode set CV*), the errors in source localization got even bigger and the errors increased with linear approximation. Again, the simple linear approximation lead to the biggest errors in source localization also in this context of a bigger database. For the ICBMny generic FEM head model, the results show a similar behavior concerning the exact position of the electrodes. If the exact correspondence is gone, lnMAG and RDM are not among the lowest anymore and source localization accuracies decreases as well. It is interesting as well that the low RDM and lnMAG with very little bias we have experienced in the section before seems to be not a result solely of the FEM head model. This implies that the correspondence of the electrode positions with source positions is essential.

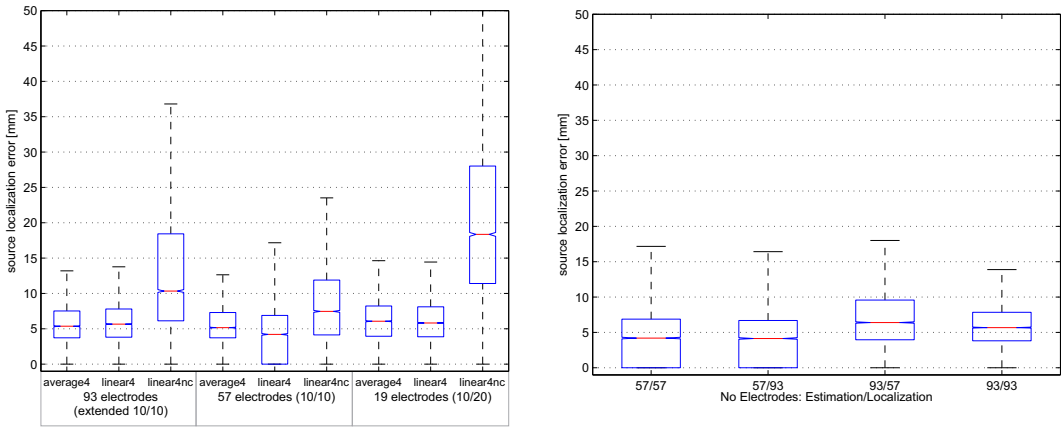

Figure 5.16: The performance of source localization in the cortex surface of different head model approximations over 4 subjects for different electrode set sizes: the electrode set including the neck area (93) increases error in the estimation of the leadfields but not in source localization.

5.4.3.2 Number and Subset of Electrodes

Investigating the effect of different electrode set sizes used for linear leadfield estimation and source localization, we find a strong influence. The scalp locations of the different sets can be found in *Figure 5.10*. As we see in *Figure 5.16*, using only the 57 electrodes from superior locations heavily improves the source localization results compared to the extended 93 positions, while only 19 electrodes seem not a good choice for approximation.

The question is, whether the reason for the dependence of the results on the electrode set is in the number of electrodes in the linear approximation or source localization or both and we investigated by using different numbers of electrodes for approximation and reconstruction.

Figure 5.16 reveals an answer to this question: both, the specific electrode set used in the linear approximation and in the source localization matters: for the linear approximation it seemed to be that excluding the extended electrode positions in neck and face helps improving the head model (compare 57/57 with 93/57) and this works also in an extrapolation to these electrode positions for source localization (57/93).

In source localization we have a different picture, as the whole set of electrodes lead to a better reconstruction (93/93), while less electrodes of the same leadfield produced worse results (93/58).

5.4.3.3 Conductivity Investigations

We will now investigate how the different head model approximations perform if they are built with wrong conductivities. In reality, the conductivities of the tissues are often

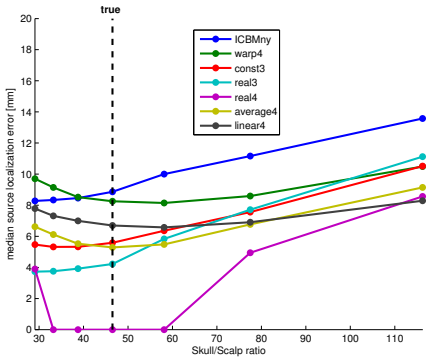

Figure 5.17: Median source localization errors for different head model approximations and conductivities: the optimal value depends on the type of model. Interestingly, for exactly the same geometry the optimum was lower for a 3-shell than a 4-shell model which corresponds to necessary higher skull-scalp ratio for proper 3-shell modeling.

not known and can vary between participants and situations.

A 4-shell BEM head model with varying scalp-to-skull conductivity ratio from 29 to 116 is used as a validation tool. Usual standard conductivity ratio assumptions lie in between 40 and 80 and this ratio is often named to be the most dominant factor in the propagation from inner fields of the head onto the scalp [Goncalves et al., 2003; Clerc et al., 2005a]. So in this case the 'ground-truth' is approximated by a 4-shell BEM head model equivalent to the one used as an approximation due to the favorable computation speed.

The results show the advantage of building an individual head model through impedance measurements if - as in reality - the individual tissue conductivities are not known. The linear non-negative approximation has the lowest variance and is, hence, the most adaptive to the individual tissue conductivity ratio of skull to scalp. If we take the same BEM head model for validation used for approximation in real4, average4 and warp4, we can see that the average lead to overall lower errors in source localization than the linear approximation.

Interestingly, 3-shell head models performed better in source localization with too high skull-to-scalp conductivity ratio (remember that the reference head models conductivities are changed and those of the approximations are kept constant). This implies a dependence of the optimal ratio on the type of head model.

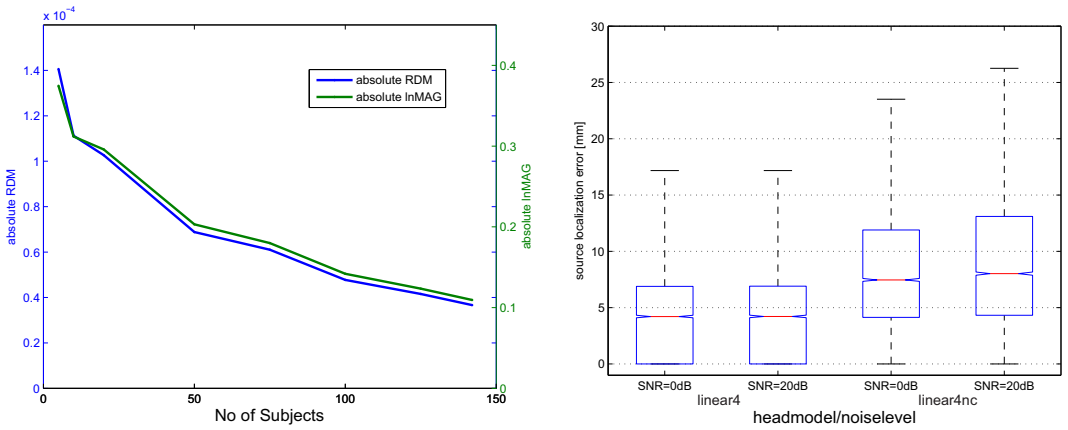

Figure 5.18: **left** Average RDM and lnMAG for non-negative linear impedance based approximation compared to the size of the database in use: we can see a smooth decrease slowly saturating at about over 100 subjects in the database. **right** Mean source localization error over relative noise level in impedance measurements: noise notably increases the error in the unconstrained (linear4uc) while not in the non-negative case (linear4).

5.4.3.4 Size of the Database & Internal Validation

There are two opposing requisites for the model: quality of reconstruction on the one hand and computational factors on the other hand. In an internal validation within the database smaller subset of subjects are randomly chosen in order to determine a useful size for the database. Each of these experiments was repeated 20 times to minimize sampling effects. The resulting leadfields are compared to the 4-shell BEM leadfield extracted from the individual MRI by calculating absolute RDM and lnMAG.

We can see that a minimum of about 125 heads leads to sufficient precision. The size of the database is not limited in general, but additional work in segmentation and head modeling per subject as well as memory size are of concern. Absolute values of RDM and lnMAG monotonically decrease in the simulations. While with the 142 subjects, theses error measures have not fully converged to a minimum value, we seem close to convergence.

5.4.3.5 Noise Effects

In order to receive a basic intuition on the effect of EIT measurement noise on the linear leadfields, EIT measurements with uncorrelated Gaussian white noise were simulated. The results in *Figure 5.18* suggest on average almost no change over subjects, although a relatively strong measurement noise corresponding to 20dB SNR is added. Only for the

linear (unconstrained) leadfields, there is a noticeable degradation in source localization performance.

5.4.4 Approximations by Impedance

This section shows why and when the average and linear leadfields work well: if the corresponding electrode locations and grid locations are perfectly aligned. Within the average and linear leadfields, a non-linearly warped electrode and source position alignment is applied. A small shift of the electrode positions strongly increases the error, because this precise alignment is lost. The electrode location is crucial for an average leadfield as well as the linear approximations.

With known electrode locations an impedance-based linear leadfield approximation of a database with corresponding anatomical locations performs comparably to the individual realistic 4-shell head model. It is robust to noise. Also, the average leadfield is a better approximation than the warp-based equivalent, a finding that is even stronger than the similar findings by [Valdés-Hernández et al., 2009]. The linear leadfield is able to adopt to individual conductivities as it shows less effect of wrong conductivity values than any other approach. This is a definite advantage as usually individual conductivities are not known and assumed to be standard although they are not expected to be.

In a more realistic setting (electrode-set CV), the performance of the average leadfield drops drastically and also the linear leadfield does not bring much of a difference. Only if we include the correctly mapped electrodes into the database, the linear approximation improves results in most cases (3 out of 4 subjects).

This also implies that a simple knowledge of the electrode positions does not necessarily help with source localization. Known electrode locations imply an estimate of where the sources lie relative to the electrodes. How the brain is folded and in particular the structure of CSF and skull is hidden implicitly in the non-linear warp, which is warped accordingly with the electrodes. In this abstract scenario, everything works perfectly, but errors are strongly underestimated in terms of the sources' real brain area compared to a setting where we do not know this warp. If we know the warp, we know the MRI and have normalized it and thus an individual head model is not difficult to obtain and possibly the better solution.

On the other hand, the linear approach seems to work if the individual head is representative for the database. This means that head models with correct electrode locations have to be in the database and the individual head must not be an outlier for the procedure to work. This could imply that for the linear leadfield approach, a huge database with a large number of heads and many different electrode positions could make it work. As our database is already in the range of Gigabytes, it will be a matter of memory rather then calculation time. In combination with the linear reconstruction, the extended (lower) electrode positions seem to increase error, while for the consecutive

source localization they improve the results. The results that a linear leadfield based only on superior electrode positions is better in later source localization can be interpreted in an effect of the head model errors of a 4-shell BEM head models in the lower head areas.

The unconstrained linear approximation is among the least reliable in almost all cases and is more affected by noise then the non-negative. An unconstrained linear interpolation might compensate too much for these systematic errors between our type of BEM and a highly detailed FEM (or a real head). This leads to artificial combinations not based so much on the geometry but on the modeling errors. The stronger effect of noise on the reconstruction quality can also be interpreted as an over-fitting towards modeling errors.

The warp based only on the outer headshape, however, is not worth its tedious construction: the average leadfield performs comparably even if exact electrode positions were unknown. The ICBMny produces source reconstruction errors on the same scale as the average leadfield, but amplitude errors might be lower in reality, which makes it favorable for applications like tCS targeting, where they are important.

Also, a different regularization term in the least squares fit could improve results. The fact that non-negative linear least squares produces sparser and better solutions than the unconstrained regular linear least squares points into this direction. The regular linear least squares is more affected by measurement noise and the choice of the database electrode sets which could be interpreted as an over-fitting problem towards model and measurements errors.

The calculation of the linear approximations is quick and takes about 90s on a 64 bit Windows 8.1 PC with Quad-Core 2.4 GHz CPU and 16 GB RAM. The average is calculated even faster and it is computation is only necessary once for all possible subjects, except if we want to incorporate the individual electrode positions.

5.5 Chapter Summary

As expected, increased realism improves the quality of head modeling and thus source localization. Whether the individual anatomy is known or not, the individual tissue conductivities, the electrode and source locations are essential factors. All can be estimated in a convenient way using impedance measurements.

As a result of the linear leadfields approach, not all common approaches seem to justify their tedious generation process. For the novel linear approach in *Section 5.4*, the generation of an individual model for electrical volume conduction of the human head through impedance measurements is successful, leading to results of similar quality as an individual head model of the same type, see *Section 5.4.3.1*. It is based on a simple linear superposition of leadfields based on the individual impedances. The linear approximation improves performance compared to an average head model and leads

to similarly low errors as the individual 4-shell head model. Compared to an FEM head model in section *Section 5.4.3.1*, this is a good estimation for the generalization performance on real heads. However, the section also shows that correspondence of electrode and source locations between database and individual are of main effect or, more precisely, the basis of linear or average leadfields to perform well.

A warped template MRI through electrode locations is also a reasonable but not necessarily the best approximation. Average leadfields are good approximations, given a good correspondence of electrode and source locations.

The more realistic the geometry for approximate or average headmodels is the better the source localization. This implies that common 3-shell average head models are not the best choice. Adding the CSF improves performance in any case since it is essential for the realism of the model as also underlined in *Chapter 2*.

For all approaches, electrode and source location correspondence are essential for the error to remain low. Also, the conductivity has to be appropriate for the type of model. The optimal conductivity depends on the model as it is dependent on the geometric errors made by the simplifications of the approach. Hence, it is not surprising that different values for the scalp/skull conductivity ratio have been found in the past, depending on the model and the method of estimation. An individual fit based on impedance measurement does not neccessarily increase performance as it might counteract geometric errors in a way favorable for scalp current injection. Most of the current flows within the scalp and inner tissue conductivity are used to minimize the scalp error without considering the error made inside the head.

The appropriate optimization methods improve the estimation process and geometric dimensionality reduction is highly successful. Combining these two approaches in the future with an alternating optimization scheme as a basis for joint shape and conductivity optimizations can prevent both processes from balancing the error made by the other. This could additionally include an electrode position estimation.

For the experimental practice, it is mainly necessary that the source and sensor locations correspond topographically, which can be either succeeded by detailed measurements or estimations of all parameters or by highly standardized protocols that guarantee their correspondence. The flexible caps or similar positioning approaches guarantee a basic correspondence but for better results, they should be at least measured and relative position of the sources in the brain estimated. Otherwise, source localization, reconstruction and the assessment of functional connectivity remain imprecise.

Chapter 6

Discussion and Outlook

In this thesis, the importance of theoretical considerations for improving application was shown. Machine learning as a statistical, data-driven approach gains widely from the inclusion of prior knowledge in particular with relatively little training data as in the case of EEG.

The details of volume conduction imply a variety of factors that have to be obeyed not only for proper modeling of the forward field of EEG but also in its analysis. Electrostatics lead to linearity in the propagation from sources to sensors, which in turn defines linear spatial filtering as optimal for the detection of distributed sources over the brain. The choice of model is important, as a minimum for appropriate source localization is the inclusion of the CSF in a 4-shell BEM head model. While, in general, the more realistic the underlying geometric representation is the better the approximation, also parameters like tissue conductivities and electrode and source locations have to be well estimated in order for the outcome to be satisfactory. Impedance measurements can be used to optimize the forward model of EEG where it is used and is beneficial to the quality of the results.

The linearity in field propagation leads to a very comfortable situation: the principle of superposition. This implies the theoretical possibility to uniquely and linearly unmix all present sources of a signal: linear spatial filtering is optimal. Blind source separation techniques only based on the statistics of the signal can help but often lead to implausible or mixed components due to the low number of independent sensors and the limited number of samples. Some unsupervised algorithms like SSD that incorporate prior knowledge on the underlying processes help to identify neural sources. Supervised approaches that incorporate the experimental design into the filter like CSP select only the relevant sources and a combination of SSD and CSP is helpful. Sequentially applying the two algorithms leads to a first selection of neural oscillations that is followed by the extraction of the relevant ones.

A main issue remains: the non-stationarity of EEG data in combination with its high autocorrelation due to spatial smearing and slow-scale temporal behavior. This non-stationarity is based on various parameters like the electrode impedance, the activity of the sweat cells, artefact levels but also the fluctuating brain state. We can find the non-stationarity directly in the PSD as the $\frac{1}{f^{\alpha}}$ structure of the whole spectrum, of sub-

bands and in the amplitudes of bandpass-filtered data or single peaks which are actually related. This leads to the fact that extracting stationarity components as with sCSP can decrease performance in a motor-imagery experiment as discriminative information is reduced. This might be partly related to discriminative artefacts, but rather also non-stationary brain signals were dismissed.

Another factor is that while it is common in the neuroscience community to talk about distinct sub-bands of the EEG spectrum - like alpha, beta and gamma - their independence is not neccessarily given. They do not describe distinct mechanisms but form an interrelated continuum of dynamic brain activity. We have seen that they can even be simple spectral harmonics of single frequency brain oscillations. This can be explained by the non-linearity in firing rate of single neurons in combination with the feedback on many levels of the neural network. Also, we should never forget that we measure predominantly the PSPs as the input to a certain area in the lower frequencies. The actual amplitude change occurs elsewhere and has a non-linear effect. This also has implications for assessing functional connectivity: a linear relationship as often assumed is very unlikely. We should rather incorporate the non-linearity into the methods of estimation and the fact that we measure the activity neural arriving from other areas.

A good analogy is that of partial and homogeneous solutions in solving differential equations like the volume conduction in chapter 2: if no or little input as an external driver (partial solution) is fed into the network, the harmonic solutions dominate the field. This homogeneous solution is found in resting state data and linked to the eigenvalues of the Laplace operator of the dispersion relation - another differential equation. It is constructed by brain connectivity in combination with transmission delays and defines the basic modes of oscillations within the brain. These are dependent on the size of the sub-network but also the properties of the connections due to the limited and variable transmission speed of axons.

We can actually set up a differential equations that incorporates both, the linear volume conduction *Equation* (2.3) and the non-linear neural transmission *Equation* (2.15):

$$\nabla \cdot (\sigma \nabla \Phi) = f(x,t) = g(\Delta \Psi(x,t)) \tag{6.1}$$

A similar approach is performed in [Nunez and Srinivasan, 2006]. If we incorporate all dependencies of the local current dipole from the neural activities at location x and time t into the function $g(\Psi(x,t))$, we have an appropriate basis for the description and estimation of functional connectivity. The key part is to include the sigmoid relationship into the dependency of Ψ on neural activity as in [Jirsa and Haken, 1997].

This harmonic non-linear dynamic system is distorted by event related activity. An event can also be memory consolidation, sleep or similar intrinsic events. This 'external' driver can be a single event but also incorporates further oscillations. Steady-state visually evoked potentials (SSVEPs) for example drive the brain to oscillate in the frequency of stimulation. Visual and auditory input is generally oscillatory in space

and time while for auditory perception the temporal and for visual the spatial domain dominate.

Due to the fact that inhibition is a rather fast local phenomenon within cortical areas, we can investigate slow large-scale and fast small-scale activity separately to some extent. The large-scale and small-scale description actually accounts for spatial and electrical amplitude scale, as larger synchrony and spatial spread leads to larger signals. Within the slow large-scale investigations, a similar distinction can be further conducted.

Let us now look at the large-scale first: if the network is disturbed, this can have various effects including the general increase in excitation and the increase in a certain frequency range. It can be spatially broad or limited to sub-networks. For the sub-network this implies a non-linear change in firing rate of which the amplitude of an oscillation's base frequency and harmonics is modulated in a non-linear way. The spectral composition of the harmonics is also influenced by background processes not directly related to the oscillation itself. As this output is relayed to other areas, it has an influence on their behavior.

The networks consist of global and local oscillations and stochastic resonances. Both actually form a continuum and I hereby propose that the faster and more local oscillations are influenced by the harmonics of the base oscillation. This implies that an amplitude change in the alpha oscillation of the Thalamus, that leads to an increase in the beta peak due to the non-linear creation of harmonics, influences more local oscillations within the cortices. The preferred frequency of each oscillation or resonance makes them lock-in to the harmonics of the driver within their range which is a basic physical phenomenon.

The fact that - apart from gamma activity - we can mainly measure low-frequency oscillations within the brain could mean that there are not many harmonics or they are actually too small and local to be measured. The latter hypothesis is supported by the basic principles described within this thesis.

Looking at the fast small-scale activity related to local inhibition, we find mainly a structuring of broadband gamma activity that bundles temporally and topographically related and inhibits unrelated activity. It thus enhances and suppresses ongoing activity and sorts it into different phases of slower oscillations depending on the cortical location. This is related to traveling waves of the slow oscillations due to the limited transmission speed within the brain. Direct long-range connections link different areas of the brain and thus activate the multimodal quality of associations.

Some researches proclaim an inhibition hypothesis that relates alpha band activity to an active suppression mechanism and not an idling mechanism as assumed earlier. This theory is mainly based on the fact that a local decrease in alpha can lead to an alpha increase in the surround in a topographical center-surround pattern within the cortex. While this is still a possible explanation, we have seen, that an increase in central alpha in combination with a general increase of excitation level non-linearly introduces a decrease in local alpha and an increase in the harmonics due to the soft-clipping effect of the non-

linearity. If alpha increases more globally and the general excitation level increases more locally, this could have a similar effect.

The brain plays its own harmonics and this is most probably related to the change of brain state that music can elicit: simple sounds or music based on natural harmonics can calm us down but can also be boring after a while. Also structured noise like the waves of an ocean can calm us down. They act like meditation where the EEG spectrum resembles that of resting state. It helps us think on larger scales and many insights that transfer knowledge from one to another domain happen while relaxing or meditating. More complex sounds or noise in particular in higher frequencies activate many different mostly highly detailed perceptions and associations.

Appendices

A χ^2 distributions

The χ^2 distribution with k degrees of freedom is defined by its probability density function $p_x(x)$:

$$X \sim \chi_k^2 : p_x(x) = \frac{1}{2^{\frac{k}{2}}\Gamma\left(\frac{k}{2}\right)} x^{\frac{k}{2}-1} e^{-\frac{x}{2}} \tag{A.1}$$

for $x > 0$ and 0 else for a random variable X that is the sum of i squared Normal distributed variables Z_i, so $X = \sum_{i=1}^{n} Z_i^2$. The first and second moments are simple: the mean $\mu_x = k$ and the variance $\sigma_x^2 = 2k$.

A.1 Generalized χ^2-distributions

Let us first introduce a general relationship between a random variable X and a linearly scaled version $X = aY$ of it [Bishop, 2006]:

$$p_y(y) = \left|\frac{d(x)}{dy}\right| p_x(x) = \left|\frac{d(ay)}{dy}\right| p_x(ay) = a p_x(ay) \tag{A.2}$$

with $a = \frac{k}{\sigma_z^2}$. The scaling factor of $\left|\frac{d(x)}{dy}\right|$ comes from the fact that we are manipulating the scale of a probability density and hence to still fulfill in particular the normalization constraint $\int_{-\infty}^{\infty} p_y(y)dy = 1$ we need this factor.

If Z is Normal distributed with zero mean and unit variance and we measure the variance $\sigma_z^2 = \frac{1}{n-1}\sum_{i=1}^{n}(Z-\mu_z)^2 = \frac{1}{n-1}\sum_{i=1}^{n} Z^2 = \frac{1}{n-1}X$, then σ_z^2 is χ^2 distributed with $k = n-1$ degrees of freedom, hence $\sigma_z^2 \sim \chi_k^2 : p_{\sigma_z^2}(\sigma_z^2) = k p_x(k\sigma_z^2)$.

If Z is Gaussian distributed with zero mean and variance σ_z^2 and we measure the variance $\hat{\sigma}_z^2 = Y = \frac{1}{n-1}\sum_{i=1}^{n} Z^2$ a simple scale transformation $y \to \frac{k}{\sigma_z^2} y =: ay$ does the trick to receive a χ^2distributed variable. In other words, $X = aY \sim \chi_k^2$ and hence:

$$p_y(y) = \frac{k}{\sigma_z^2} \frac{1}{2^{\frac{k}{2}}\Gamma\left(\frac{k}{2}\right)} \left(\frac{ky}{\sigma_z^2}\right)^{\frac{k}{2}-1} e^{-\frac{ky}{2\sigma_z^2}}$$

$$p_y(y) = \left(\frac{k}{\sigma_z^2}\right)^{\frac{k}{2}} \frac{1}{2^{\frac{k}{2}}\Gamma\left(\frac{k}{2}\right)} y^{\frac{k}{2}-1} e^{-\frac{ky}{2\sigma_z^2}} \tag{A.3}$$

The mean μ_y and variance σ_y^2 are:

$$\mu_y = \int_0^\infty p_y(y) y dy = \int_0^\infty a p_x(ay) y dy = a \int_0^\infty p_x(x) \frac{x}{a} \frac{dx}{a} = \frac{k}{a} = \sigma_z^2$$

$$\sigma_y^2 = \int_0^\infty p_y(y)(y - \frac{k}{a})^2 dy = \int_0^\infty a p_x(ay) \left(y - \frac{k}{a}\right)^2 dy = a \int_0^\infty p_x(x) \left(\frac{x}{a} - \frac{k}{a}\right)^2 \frac{dx}{a}$$

$$= \frac{1}{a^2} \int_0^\infty p_x(x) \left(x - k\right)^2 dx = \frac{2k}{a^2} = 2\frac{\sigma_z^4}{k}$$

To fit any measured distribution to a generalized χ^2 distribution by its mean and variance, a virtual degree of freedom can be developed by measuring the variance $\hat{\sigma}_y^2$ and the mean $\hat{\mu}_y$ of the variance data and then choosing $\hat{a}$ and $\hat{k}$ accordingly. This leads to:

$$\hat{k} = 2 * \frac{\hat{\mu}_y}{\hat{\sigma}_y^2}$$

$$\hat{a} = \frac{\hat{k}}{\hat{\mu}_y} = \frac{2}{\hat{\sigma}_y^2}$$

A.2 Optimal χ^2 classifier

If we have variance data x incorporating two classes with generalized χ^2 distributions with parameters $k_i^{(c)}$ and $d_i = \mu_{\sigma_z^2}$, we can calculate the class-wise distributions for each channel by (compare *Equation* (A.3)):

$$p_{x_i}(x_i) = \left(\frac{k_i^{(c)}}{d_i}\right)^{\frac{k_i^{(c)}}{2}} \frac{1}{2^{\frac{k_i^{(c)}}{2}}\Gamma\left(\frac{k_i^{(c)}}{2}\right)} x_i^{\frac{k_i^{(c)}}{2}-1} e^{-\frac{k_i^{(c)} x_i}{2d_i}}$$

In the case of two-class CSP, the two classes are separated in an orthogonal manner by the algorithm, leading to two different distributions with the d_i for class 1 and $1 - d_i$ for class two. The resulting distribution for each channel over both classes is the simple sum of the distributions of each class.

We can now set up the modified χ^2 function for both classes $p_y(y_1, y_2, ...y_n)$ with $y_i = \frac{k_i^{(1)}}{d_i} x_i$ and $p_z(z_1, z_2, ...z_n)$ with $z_i = \frac{k_i^{(2)}}{1-d_i} x_i$

The optimal separating hyperfunction between the two classes is where the probability of belonging to class 1 is equal to the probability of belonging to class 2, hence, where the difference is zero:

$$p_y(y_1, y_2, ...y_n) = p_z(z_1, z_2, ...z_n)$$
$$\prod_{i=1}^{n} \frac{k_i^{(1)}}{d_i} \frac{1}{2^{\frac{k_i^{(1)}}{2}} \Gamma\left(\frac{k_i^{(1)}}{2}\right)} y_i^{\frac{k_i^{(1)}}{2}-1} e^{-\frac{y_i}{2}} = \prod_{i=1}^{n} \frac{k_i^{(2)}}{1-d_i} \frac{1}{2^{\frac{k_i^{(2)}}{2}} \Gamma\left(\frac{k_i^{(2)}}{2}\right)} z_i^{\frac{k_i^{(2)}}{2}-1} e^{-\frac{z_i}{2}}$$

This is equivalent to setting the likelihood ratio to 1 from a Bayesian perspective of the discrimination task. Logarithmizing yields:

$$\log p_y(y_1, y_2, ...y_n) = \sum_{i=1}^{n} \left[\log\left(\frac{k_i^{(1)}}{d_i}\right) - \left(\frac{k_i^{(1)}}{2}\right) \log 2 - \log \Gamma\left(\frac{k_i^{(1)}}{2}\right) + \left(\frac{k_i^{(1)}}{2} - 1\right) \log y_i - \frac{y_i}{2} \right]$$

and so:

$$\log p_y(y_1, y_2, ...y_n) = \log p_z(z_1, z_2, ...z_n)$$
$$\sum_{i=1}^{n} \left[\log\left(\frac{k_i^{(1)}(1-d_i)}{k_i^{(2)} d_i}\right) - \frac{k_i^{(1)} - k_i^{(2)}}{2} \log 2 + \log\left(\frac{\Gamma\left(\frac{k_i^{(2)}}{2}\right)}{\Gamma\left(\frac{k_i^{(1)}}{2}\right)}\right) + ... \right.$$
$$\left. \left(\frac{k_i^{(1)}}{2} - 1\right) \log y_i - \left(\frac{k_i^{(2)}}{2} - 1\right) \log z_i - \frac{y_i}{2} + \frac{z}{2} \right] = 0$$

Inserting $y_i = \frac{k_i^{(1)}}{d_i} x_i$ and $z_i = \frac{k_i^{(2)}}{1-d_i} x_i$ results in:

$$\sum_{i=1}^{n} \left[\log\left(\frac{k_i^{(1)}(1-d_i)}{k_i^{(2)} d_i}\right) + \frac{k_i^{(2)} - k_i^{(1)}}{2} \log 2 + \log\left(\frac{\Gamma\left(\frac{k_i^{(2)}}{2}\right)}{\Gamma\left(\frac{k_i^{(1)}}{2}\right)}\right) + ... \right.$$
$$\left. \log\left(\frac{\left(\frac{k_i^{(1)}}{d_i}\right)^{\frac{k_i^{(1)}}{2}-1}}{\left(\frac{k_i^{(2)}}{1-d_i}\right)^{\frac{k_i^{(2)}}{2}-1}}\right) + \log\left(\frac{x_i^{\frac{k_i^{(1)}}{2}-1}}{x_i^{\frac{k_i^{(2)}}{2}-1}}\right) - \frac{x_i k_i^{(1)}}{2d_i} + \frac{x_i k_i^{(2)}}{2(1-d_i)} \right] = 0$$

$$\sum_{i=1}^{n}\left[\log\left(\frac{k_i^{(1)}\left(1-d_i\right)}{k_i^{(2)}d_i}\right)+\frac{k_i^{(2)}-k_i^{(1)}}{2}\log 2+\log\left(\frac{\Gamma\left(\frac{k_i^{(2)}}{2}\right)}{\Gamma\left(\frac{k_i^{(1)}}{2}\right)}\right)+\ldots\right.$$
$$\left.\log\left(\frac{\left(\frac{k_i^{(1)}}{d_i}\right)^{\frac{k_i^{(1)}}{2}-1}}{\left(\frac{k_i^{(2)}}{1-d_i}\right)^{\frac{k_i^{(2)}}{2}-1}}\right)+\frac{k_i^{(1)}-k_i^{(2)}}{2}\log x_i+\frac{1}{2}\left(\frac{k_i^{(2)}}{1-d_i}-\frac{k_i^{(1)}}{d_i}\right)x_i\right]=0 \tag{A.4}$$

which resembles a non-linear (probabilistic) classifier:

$$y=\sum_{i=1}^{n}\left(a_i+b_i\log x_i+c_i x_i\right) \tag{A.5}$$

If the (virtual) degrees of freedom of both classes are equal $k_i^{(1)}=k_i^{(2)}=k_i$, this simplifies to:

$$\frac{1}{2}\sum_{i=1}^{n}\left[k_i\left(\frac{1}{1-d_i}-\frac{1}{d_i}\right)x_i\right]+\sum_{i=1}^{n}\left[\left(\frac{k_i}{2}-1\right)\log\frac{1-d_i}{d_i}\right]=0$$

This actually resembles a linear classifier in the variance space:

$$y=\mathbf{w}^T\mathbf{x}+b \tag{A.6}$$

with $w_i=\frac{k_i}{2}\left(\frac{1}{1-d_i}-\frac{1}{d_i}\right)$ and $b=\sum_{i=1}^{n}\left(\frac{k_i}{2}-1\right)\log\frac{1-d_i}{d_i}$.

We can also linearize *Equation* (A.5) in any point $\mathbf{q}$ using its Taylor series to receive a linear classifier:

$$y\approx f(\mathbf{q})+\nabla f(\mathbf{q})\cdot(\mathbf{x}-\mathbf{q})=:\mathbf{w}^T\mathbf{x}+b \tag{A.7}$$

using $\frac{df(\mathbf{q})}{dq_i}=\left(\frac{b_i}{q_i}+c_i\right)$.

The classifier output y in all cases directly relates to the probabilities. y is the difference in logarithmic probability between the two classes. This can be translated into a likelihood ratio t by

$$t=e^y.$$

B Conductivity Derivatives for the 4-shell Boundary Element Method

Here, it is convenient to first compute the derivatives w. r. t. $\mathbf{A}$, and then to build the inverse through the rules:

$$\frac{d\mathbf{A}^{-1}}{d\theta_i} = -\mathbf{A}^{-1}\frac{d\mathbf{A}}{d\theta_i}\mathbf{A}^{-1} \tag{B.1}$$

and

$$\frac{d^2\mathbf{A}^{-1}}{d\theta_i d\theta_j} = -\frac{d\mathbf{A^{-1}}}{d\theta_i}\frac{d\mathbf{A}}{d\theta_j}\mathbf{A}^{-1} - \mathbf{A}^{-1}\frac{d^2\mathbf{A}}{d\theta_i d\theta_j}\mathbf{A}^{-1} - \frac{d\mathbf{A^{-1}}}{d\theta_j}\frac{d\mathbf{A}}{d\theta_i}\mathbf{A}^{-1}.$$

Note the advantage that the symmetric BEM leads to a symmetric matrix which is in general regular or close to it.

Taking the first derivative of equation (5.2) towards θ_i leads to:

$$\frac{d\hat{\mathbf{v}}\left(\theta_i\right)}{d\theta_i} = \mathbf{W}\left(\frac{d\mathbf{A}^{-1}}{d\theta_i}\mathbf{b} + \mathbf{A}^{-1}\frac{d\mathbf{b}}{d\theta_i}\right) \tag{B.2}$$

and the second is:

$$\frac{d^2\hat{\mathbf{v}}\left(\theta_i\right)}{d\theta_i d\theta_j} = \mathbf{W}\left(\frac{d^2\mathbf{A}^{-1}}{d\theta_i d\theta_j}\mathbf{b} + \frac{d\mathbf{A}^{-1}}{d\theta_i}\frac{d\mathbf{b}}{d\theta_j} + \frac{d\mathbf{A}^{-1}}{d\theta_j}\frac{d\mathbf{b}}{d\theta_i} + \mathbf{A}^{-1}\frac{d^2\mathbf{b}}{d\theta_i d\theta_j}\right). \tag{B.3}$$

We based our BEM model on the set of linear equations defined in [Clerc et al., 2005b]. The symmetric BEM models interface potentials and currents simultaneously. For the system

$$\mathbf{A}_\theta[\mathbf{x}_1, \mathbf{y}_1, \mathbf{x}_2, \mathbf{y}_2, \mathbf{x}_3, \mathbf{y}_3, \mathbf{x}_4]^\top = \mathbf{b}\ , \tag{B.4}$$

where the vectors $\mathbf{x_k}, k = 1, \ldots, 4$ correspond to the potential on the interfaces S_k of the different compartments (scalp, skull, CSF and brain), and where the vectors $\mathbf{y_k}$ correspond to its derivative (normal currents), we get

$$\mathbf{A}\left(\theta\right) = \begin{bmatrix}
\left(\sigma_1+\sigma_2\right)\mathbf{N}_{11} & -2\mathbf{D}^*_{11} & -\sigma_2\mathbf{N}_{12} & \mathbf{D}^*_{12} & 0 & 0 & 0 \\
-2\mathbf{D}_{11} & \left(\sigma_1^{-1}+\sigma_2^{-1}\right)\mathbf{S}_{11} & \mathbf{D}_{12} & -\sigma_2^{-1}\mathbf{S}_{12} & 0 & 0 & 0 \\
-\sigma_2\mathbf{N}_{21} & \mathbf{D}^*_{21} & \left(\sigma_2+\sigma_3\right)\mathbf{N}_{22} & -2\mathbf{D}^*_{22} & -\sigma_3\mathbf{N}_{23} & \mathbf{D}^*_{23} & 0 \\
\mathbf{D}_{21} & -\sigma_2^{-1}\mathbf{S}_{21} & -2\mathbf{D}_{22} & \left(\sigma_2^{-1}+\sigma_3^{-1}\right)\mathbf{S}_{22} & \mathbf{D}_{23} & -\sigma_3^{-1}\mathbf{S}_{23} & 0 \\
0 & 0 & -\sigma_3\mathbf{N}_{23} & \mathbf{D}^*_{32} & \left(\sigma_2+\sigma_3\right)\mathbf{N}_{33} & -2\mathbf{D}^*_{33} & -\sigma_4\mathbf{N}_{34} \\
0 & 0 & \mathbf{D}_{23} & -\sigma_3^{-1}\mathbf{S}_{32} & -2\mathbf{D}_{33} & \left(\sigma_3^{-1}+\sigma_4^{-1}\right)\mathbf{S}_3 & \mathbf{D}_{34} \\
0 & 0 & 0 & 0 & -\sigma_3\mathbf{N}_{34} & \mathbf{D}^*_{43} & \sigma_4\mathbf{N}_{44}
\end{bmatrix},$$

while for the EIT case we obtain

$$\mathbf{b}\left(\theta\right) = \left[0, 0, 0, 0, -\mathbf{D}^*_{34}\mathbf{z}, \sigma_4^{-1}\mathbf{S}_{34}\mathbf{z}, \left(-\frac{1}{2}\mathbf{I}_{44} + \mathbf{D}^*_{44}\right)\mathbf{z}\right]^\top\ . \tag{B.5}$$

Here, $\mathbf{z}$ is the discretization of the normal currents on the scalp so assembling our injected currents and generally $\theta = (\sigma_1, \sigma_2, \sigma_3, \sigma_4, \mathcal{S}_1, \mathcal{S}_2, \mathcal{S}_3, \mathcal{S}_4)$ involving the conductivities σ_i

and the shapes of the interfaces $\mathcal{S}_i$. The matrix elements $\mathbf{N}_{ij}$,$\mathbf{S}_{ij}$,$\mathbf{D}_{ij}$ and $\mathbf{D}^*_{ij}$ are block matrices assembling discretized versions of the integral operators of classical (Newtonian) potential theory between layers i and j. $\mathbf{I}_{ij}$ is an identity operator.

From this we find the block-elements of the Jacobian towards the conductivities σ (keeping the geometry constant) as:

$$\frac{d\mathbf{A}}{d\sigma_1} = \begin{bmatrix} \mathbf{N}_{11} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -\sigma_1^{-2}\mathbf{S}_{11} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \quad \frac{d\mathbf{A}}{d\sigma_2} = \begin{bmatrix} \mathbf{N}_{11} & 0 & -\mathbf{N}_{12} & 0 & 0 & 0 & 0 \\ 0 & -\sigma_2^{-2}\mathbf{S}_{11} & 0 & \sigma_2^{-2}\mathbf{S}_{12} & 0 & 0 & 0 \\ -\mathbf{N}_{21} & 0 & \mathbf{N}_{22} & 0 & 0 & 0 & 0 \\ 0 & \sigma_2^{-2}\mathbf{S}_{21} & 0 & -\sigma_2^{-2}\mathbf{S}_{22} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

$$\frac{d\mathbf{A}}{d\sigma_3} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \mathbf{N}_{22} & 0 & -\mathbf{N}_{23} & 0 & 0 \\ 0 & 0 & 0 & -\sigma_3^{-2}\mathbf{S}_{22} & 0 & \sigma_3^{-2}\mathbf{S}_{23} & 0 \\ 0 & 0 & -\mathbf{N}_{32} & 0 & \mathbf{N}_{33} & 0 & 0 \\ 0 & 0 & 0 & \sigma_3^{-2}\mathbf{S}_{32} & 0 & -\sigma_3^{-2}\mathbf{S}_{33} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \quad \frac{d\mathbf{A}}{d\sigma_4} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \mathbf{N}_{33} & 0 & -\mathbf{N}_{34} \\ 0 & 0 & 0 & 0 & 0 & -\sigma_4^{-2}\mathbf{S}_{33} & 0 \\ 0 & 0 & 0 & 0 & -\mathbf{N}_{43} & 0 & \mathbf{N}_{44} \end{bmatrix},$$

whereas all $\frac{d\mathbf{b}}{d\sigma_i}$ are zero except for :

$$\frac{d\mathbf{b}}{d\sigma_4} = \left[0, 0, 0, 0, 0, -\sigma_4^{-2}\mathbf{S}_{34}z, 0\right]^\top .$$

For the derivatives necessary for the Hessian blocks, we get:

$$\frac{d^2\mathbf{A}}{d\sigma_1^2} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2\sigma_1^{-3}\mathbf{S}_{11} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \quad \frac{d^2\mathbf{A}}{d\sigma_2^2} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2\sigma_2^{-3}\mathbf{S}_{11} & 0 & -2\sigma_2^{-3}\mathbf{S}_{12} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -2\sigma_2^{-3}\mathbf{S}_{21} & 0 & 2\sigma_2^{-3}\mathbf{S}_{22} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

$$\frac{d^2\mathbf{A}}{d\sigma_3^2} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2\sigma_3^{-3}\mathbf{S}_{22} & 0 & -2\sigma_3^{-3}\mathbf{S}_{23} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -2\sigma_3^{-3}\mathbf{S}_{32} & 0 & 2\sigma_3^{-3}\mathbf{S}_{33} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \quad \frac{d^2\mathbf{A}}{d\sigma_4^2} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2\sigma_4^{-3}\mathbf{S}_{33} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

and all off-diagonals $\frac{d^2\mathbf{A}}{d\sigma_i d\sigma_j}$ are zero.

All $\frac{d^2\mathbf{b}}{d\sigma_i\sigma_j}$ are also zero, except for :

$$\frac{d^2\mathbf{b}}{d\sigma_4^2} = \left[0, 0, 0, 0, 0, 2\sigma_4^{-3}\mathbf{S}_{34}z, 0\right]^\top .$$

Note also that the values of the block matrix elements in A are known and simple multiplications by functions of the conductivity values σ_i lead to the resulting values of the derivatives.

C Error Measures

In order to estimate how well the approaches work in realistic settings, we need to take measures to compare them to each other. Within the experiments, we used the following measure to evaluate the performance of the approaches:

Logarithmic Magnitude Difference Measure To evaluate the quality of the obtained leadfield magnitudes, we employ the logarithmic magnitude difference measure (lnMAG) [Güllmar et al., 2010].

$$\text{lnMAG} = \log \frac{\left\|\hat{L}\right\|_2}{\|L\|_2}$$

The absolute magnitudes are not important in EEG source localization, as only the location of the best fitting dipole is of interest. In electrical current stimulation, however, the strength of the field induced by a current injection highly matters. This is represented by the lnMAG which we employ separately for every single source location here (the norm is over scalp potentials). We used the natural logarithm. An lnMAG of 0 means perfect amplitude correspondence.

Relative Difference Measure The relative difference measure (RDM) [Meijs et al., 1989] is often used as a measure to compare leadfields. It focuses on the scalp patterns created by leadfields regardless of amplitude. These pattern are important in source localization and low RDMs are related to better source localization [Dannhauer et al., 2011]. We calculated it for every source location and sensor combination i and looked at the distribution:

$$\text{RDM} = \frac{L_i}{\||L\|\,|_2} - \frac{\hat{L}_i}{\left\||\hat{L}\right\|\,|_2}$$

where L_i is the leadfield amplitude at location and sensor combination i and $\hat{L}_i$ the corresponding estimate. $\left\||\hat{L}\right\|\,|_2$ and $\left\||\hat{L}\right\|\,|_2$ are the L_2-norms over all source and sensor locations. An RDM of 0 means no error made.

source localization Error For the source localization error E_{loc}, we use the Euclidean distance between the estimated and the simulated source location within the individual head. We used the subspace based linear-least squares dipole fitting method MUSIC algorithm [Mosher et al., 1992] to estimate the sources and calculated the error for every source i by:

$$E_{\text{loc}} = \sqrt{\||r_i - \hat{r}_i\|\,|_2}$$

where $\hat{r}i$ is the source location estimated by the MUSIC algorithm and r_i is the initially simulated source location. A source localization error E_{loc} of 0 means perfect detection.

D Technical Details on Anatomical parameters

D.1 Sensor Positions

The basis of the electrode morph were sets of 93 electrode that were placed manually under equal policies on the 4 heads we used for validation. These coordinates were then warped non-linearly to the template head and then again to every single head of the database. This lead to 4 leadfields per subject in the approximation database. The correspondence of electrode locations between subjects have an important influence on the source localization accuracies. As can be seen in *Figure 5.10* on the left, the positions of the electrode sets varied depending on which of the evaluation subjects they originated from, even though they were originally all set in standard 10/10 locations [Chatrian et al., 1985] on the original (evaluation subject's) head. Subject 3 had shifted positions towards the neck and Subject 4 was slightly shifted towards the left front, compared to Subject 1 and 2 which were closer to each other and rather in standard positions. Note again that *Figure 5.10* is displayed on a database head. On the individual subject head, the electrodes were in the standard 10/10 positions and this distortion stemmed from the non-linear warp. The median difference between the electrode position for each channel was $8.8mm$ between the sets. We will discuss the effect of knowing or not knowing the exact map by comparing the results of a leave-one-out cross-validation over electrode sets, of all sets simultaneously in the training data and the directly corresponding set for each validation head only.

D.2 Source Simulation

At first, 149 locations were excluded because they ended up out of the brain structure for some of the subjects in the database. This is due to the error made by the coarse and smoothed approximation of the surface shapes we have used: Even within subjects the surfaces do not fully correspond to each other for discretizations extracted from different MRIs of the same person. Excluding dipoles out of this surface minimizes the effect of these modeling errors.

We simulated an EEG leadfield for the remaining 3841 sources and the 93 electrodes described in the sensors *Section D.1* with the FEM head models obtained from the 4 external MRIs of *Section 5.4.2.2*. This pseudo EEG will then be compared to the approximation solely based on impedances and to others approaches through the error measures and a MUSIC based source localization described in Appendix C.

Equivalently to the BEM leadfields of the database, the resulting FEM leadfields were referenced to Iz, which was then excluded from the data. The source locations were the set on the gray-white border orthogonal to it's surface, which gave us leadfields with 92x3841 dimensions.

The restriction was based on distance from the origin of the head model in the template mid-brain ($|r| > 39mm$), and a y/z-plane below the temporal lobe($y > -12mm$,$z > -16mm$,) where x is from the left towards the right ear, y from the center of x towards the nose and z the orthogonal direction to the former two towards superior positions).

D.3 EIT Simulation

The following procedure was done to simulate the scalp potentials produced by external current injection through the electrodes: All possible pairs of the electrode set were combined by setting a positive current of $100\mu A$ at one electrode and a negative return current of the same value at another. Because the whole model was based on a linear current-voltage relationship (Ohm's law), the current amplitude actually introduces relative effects and the absolute value is not important for the simulations.

Theoretically, we could measure voltage measurements also at the injecting electrodes, if we incorporated the electrode (contact) impedance. This value is very unsure, variable and adds only little information, as we have strongly statistical dependent sensors in proximity. Therefore, to minimize the effect of electrode impedances, currents are only injected on a small number of electrodes while resulting scalp potentials are measured at all remaining channels [Holder, 2005].

Consequently, the voltages on the injecting electrodes were excluded for the measurements. This is done for the database as well as for the external MRIs. For the external MRIs we used the same head model as for the EEG simulations. The resulting EIT gain matrices were referenced to Iz, leading to a dimensionality of 92x92.

For a basic estimation of the influence of noise, uncorrelated Gaussian noise was added to the simulated measurement amplitudes in an additional analysis. In the process of obtaining real EIT measurements there are two main sources of noise: the current injection and the voltage measurements. The produced scalp voltage is luckily in the range of mV so measuring with EEG amplifiers which are designed to work in the μV range results in comparably small measurement noise. The noise in the scalp current depends on the current synthesis and the 2 different error sources are superimposed due to the linear relationship.

We have simulated different noise levels including a much stronger noise level of Gaussian noise on the measurements than is to be expected leading to an SNR of 20dB (10:1) to evaluate the maximally expected influence. Expected is around 60dB (see *Section 3.1.2.4*).

Bibliography

Acar, Z. A., Acar, C. E., and Makeig, S. (2016). Simultaneous head tissue conductivity and EEG source location estimation. *NeuroImage*, 124, Part A:168 – 180.

Acar, Z. A. and Makeig, S. (2010). Neuroelectromagnetic forward head modeling toolbox. *Journal of Neuroscience Methods*, 190(2):258–70.

Acar, Z. A. and Makeig, S. (2013). Effects of forward model errors on EEG source localization. *Brain topography*, 26(3):378–396.

Adali, T., Anderson, M., and Fu, G.-S. (2014). Diversity in independent component and vector analyses: Identifiability, algorithms, and applications in medical imaging. *IEEE Signal Processing Magazine*, 31(3):18–33.

Akalin-Acar, Z. and Gençer, N. G. (2004). An advanced boundary element method (BEM) implementation for the forward problem of electromagnetic source imaging. *Physics in Medicine & Biology*, 49(21):5011.

Aristovich, K. Y., Packham, B. C., Koo, H., dos Santos, G. S., McEvoy, A., and Holder, D. S. (2016). Imaging fast electrical activity in the brain with electrical impedance tomography. *NeuroImage*, 124:204–213.

Ashburner, J. and Friston, K. (2005). Unified segmentation. *NeuroImage*, 26:839–851.

Atasoy, S., Deco, G., Kringelbach, M. L., and Pearson, J. (2018). Harmonic brain modes: a unifying framework for linking space and time in brain dynamics. *The Neuroscientist*, 24(3):277–293.

Atasoy, S., Donnelly, I., and Pearson, J. (2016). Human brain networks function in connectome-specific harmonic waves. *Nature communications*, 7:10340.

Baillet, S., Mosher, J., and Leahy, R. (2001). Electromagnetic brain mapping. *IEEE Signal Processing Mag.*, 18:14–30.

Bair, W. and Koch, C. (1996). Temporal precision of spike trains in extrastriate cortex of the behaving macaque monkey. *Neural computation*, 8(6):1185–1202.

Bair, W., Koch, C., Newsome, W., and Britten, K. (1994). Power spectrum analysis of bursting cells in area MT in the behaving monkey. *Journal of Neuroscience*, 14(5):2870–2892.

Baker, S., Kilner, J., Pinches, E., and Lemon, R. (1999). The role of synchrony and oscillations in the motor output. *Experimental brain research*, 128(1-2):109–117.

Baker, S. N., Gabriel, C., and Lemon, R. N. (2003). Eeg oscillations at 600 hz are macroscopic markers for cortical spike bursts. *The Journal of Physiology*, 550(2):529–534.

Barber, D. and Brown, B. (1984). Applied potential tomography. *Journal of Physics E: Scientific Instruments*, 17(9):723.

Basser, P. J. (1995). Inferring microstructural features and the physiological state of tissues from diffusion-weighted images. *NMR in Biomedicine*, 8(7):333–344.

Baysal, U. and Şengül, G. (2010). Single camera photogrammetry system for EEG electrode identification and localization. *Annals of Biomedical Engineering*, 38(4):1539–1547.

Bédard, C. and Destexhe, A. (2008). A modified cable formalism for modeling neuronal membranes at high frequencies. *Biophysical journal*, 94(4):1133–1143.

Bédard, C., Kröger, H., and Destexhe, A. (2004). Modeling extracellular field potentials and the frequency-filtering properties of extracellular space. *Biophysical journal*, 86(3):1829–1842.

Bell, A. J. and Sejnowski, T. J. (1995). An information-maximization approach to blind separation and blind deconvolution. *Neural computation*, 7(6):1129–1159.

Belliveau, J., Kennedy, D., McKinstry, R., Buchbinder, B., Weisskoff, R., Cohen, M., Vevea, J., Brady, T., and Rosen, B. (1991). Functional mapping of the human visual cortex by magnetic resonance imaging. *Science*, 254(5032):716–719.

Berger, H. (1929). Über das Elektroenkephalogramm des Menschen. *Archiv für Psychiatrie und Nervenkrankheiten*, 87(1):527–570.

Berger, H. (1931). Über das Elektrenkephalogramm des Menschen. *Archiv für Psychiatrie und Nervenkrankheiten*, 94(1):16–60.

Bishop, C. M. (2006). *Pattern recognition and machine learning*. springer.

Blankertz, B., Kawanabe, M., Tomioka, R., Hohlefeld, F., Nikulin, V., and Müller, K.-R. (2008a). Invariant common spatial patterns: Alleviating nonstationarities in brain-computer interfacing. In Platt, J., Koller, D., Singer, Y., and Roweis, S., editors, *Advances in Neural Information Processing Systems 20*, pages 113–120, Cambridge, MA. MIT Press.

Blankertz, B., Lemm, S., Treder, M. S., Haufe, S., and Müller, K.-R. (2011). Single-trial analysis and classification of ERP components – a tutorial. *NeuroImage*, 56:814–825.

Blankertz, B., Sannelli, C., Halder, S., Hammer, E. M., Kübler, A., Müller, K.-R., Curio, G., and Dickhaus, T. (2010). Neurophysiological predictor of SMR-based BCI performance. *NeuroImage*, 51(4):1303–1309.

Blankertz, B., Tomioka, R., Lemm, S., Kawanabe, M., and Müller, K.-R. (2008b). Optimizing spatial filters for robust EEG single-trial analysis. *IEEE Signal Process Mag*, 25(1):41–56.

Blanz, V. and Vetter, T. (1999). A morphable model for the synthesis of 3d faces. In *Proceedings of the 26th Annual Conference on Computer Graphics and Interactive Techniques*, SIGGRAPH '99, pages 187–194, New York, NY, USA. ACM Press/Addison-Wesley Publishing Co.

Böcker, K. B., Brunia, C. H., and Cluitmans, P. J. (1994). A spatio-temporal dipole model of the readiness potential in humans. I. finger movement. *Electroencephalography and Clinical Neurophysiology*, 91(4):275–285.

Brun, A. and Englund, E. (1986). A white matter disorder in dementia of the alzheimer type: A pathoanatomical study. *Annals of Neurology*, 19(3):253–262.

Buzsaki, G. (2006). *Rhythms of the Brain*. Oxford University Press.

Buzsáki, G. (2019). *The brain from inside out*. Oxford University Press, New York, USA.

Buzsáki, G., Anastassiou, C. A., and Koch, C. (2012). The origin of extracellular fields and currents - EEG, ECoG, LFP and spikes. *Nature Reviews Neuroscience*, 13:407–420.

Cardin, J. A., Carlén, M., Meletis, K., Knoblich, U., Zhang, F., Deisseroth, K., Tsai, L.-H., and Moore, C. I. (2009). Driving fast-spiking cells induces gamma rhythm and controls sensory responses. *Nature*, 459(7247):663–667.

Carlqvist, H., Nikulin, V. V., Strömberg, J.-O., and Brismar, T. (2005). Amplitude and phase relationship between alpha and beta oscillations in the human electroencephalogram. *Medical and Biological Engineering and Computing*, 43(5):599–607.

Cash, S. and Yuste, R. (1999). Linear summation of excitatory inputs by CA1 pyramidal neurons. *Neuron*, 22(2):383–394.

Chatrian, G. E., Lettich, E., and Nelson, P. L. (1985). Ten percent electrode system for topographic studies of spontaneous and evoked EEG activities. *American Journal of EEG Technology*, 25(2):83–92.

Clerc, M., Adde, G., Kybic, J., Papadopoulo, T., and Badier, J. (2005a). In vivo conductivity estimation with symmetric boundary elements. *IJBEM*, 7:307–310.

Clerc, M., Badier, J., Adde, G., Kybic, J., and Papadopoulo, T. (2005b). Boundary element formulation for electrical impedance tomography. *ESAIM: Proceedings*, 14:63–71.

Cohen, D. (1968). Magnetoencephalography: evidence of magnetic fields produced by alpha-rhythm currents. *Science*, 161(3843):784–786.

Cole, S. and Voytek, B. (2019). Cycle-by-cycle analysis of neural oscillations. *Journal of neurophysiology*, 122(2):849–861.

Coleman, T. F. and Li, Y. (1996). An interior trust region approach for nonlinear minimization subject to bounds. *SIAM Journal on optimization*, 6(2):418–445.

Colgin, L. L. (2016). Rhythms of the hippocampal network. *Nature Reviews Neuroscience*, 17(4):239.

Congedo, M., Barachant, A., and Bhatia, R. (2017). Riemannian geometry for EEG-based brain-computer interfaces; a primer and a review. *Brain-Computer Interfaces*, 4(3):155–174.

Connor, J. and Stevens, C. (1971). Inward and delayed outward membrane currents in isolated neural somata under voltage clamp. *The Journal of Physiology*, 213(1):1–19.

Contreras, D. and Steriade, M. (1995). Cellular basis of EEG slow rhythms: a study of dynamic corticothalamic relationships. *Journal of Neuroscience*, 15(1):604–622.

Curio, G., Mackert, B.-M., Burghoff, M., Koetitz, R., Abraham-Fuchs, K., and Härer, W. (1994). Localization of evoked neuromagnetic 600 hz activity in the cerebral somatosensory system. *Electroencephalography and Clinical Neurophysiology*, 91(6):483 – 487.

Damoiseaux, J. S., Rombouts, S., Barkhof, F., Scheltens, P., Stam, C. J., Smith, S. M., and Beckmann, C. F. (2006). Consistent resting-state networks across healthy subjects. *Proceedings of the national academy of sciences*, 103(37):13848–13853.

Dannhauer, M., Lanfer, B., Wolters, C., and Knösche, T. (2011). Modeling of the human skull in EEG source analysis. *Human Brain Mapping*, pages 1383–1399.

Darvas, F., Ermer, J., Mosher, J., and Leahy, R. (2006). Generic head models for atlas-based EEG source analysis. *Human Brain Mapping*, 27(2):129–143.

Datta, A., Zhou, X., Su, Y., Parra, L., and Bikson, M. (2013). Validation of finite element model of transcranial electrical stimulation using scalp potentials: implications for clinical dose. *Journal of Neural Engineering*, 10(3):036018.

David, O. and Friston, K. J. (2003). A neural mass model for MEG/EEG:: coupling and neuronal dynamics. *NeuroImage*, 20(3):1743–1755.

Dayan, P. and Abbott, L. F. (2001). *Theoretical neuroscience: computational and mathematical modeling of neural systems*. MIT press.

De Luca, M., Beckmann, C. F., De Stefano, N., Matthews, P. M., and Smith, S. M. (2006). fMRI resting state networks define distinct modes of long-distance interactions in the human brain. *Neuroimage*, 29(4):1359–1367.

Deco, G., Jirsa, V. K., and McIntosh, A. R. (2011). Emerging concepts for the dynamical organization of resting-state activity in the brain. *Nature Reviews Neuroscience*, 12(1):43–56.

Deco, G., Jirsa, V. K., Robinson, P. A., Breakspear, M., and Friston, K. (2008). The dynamic brain: from spiking neurons to neural masses and cortical fields. *PLoS computational biology*, 4(8).

Dickson, C. T., Magistretti, J., Shalinsky, M. H., Fransén, E., Hasselmo, M. E., and Alonso, A. (2000). Properties and role of I_h in the pacing of subthreshold oscillations in entorhinal cortex layer II neurons. *Journal of Neurophysiology*, 83(5):2562–2579.

Draguhn, A., Traub, R., Schmitz, D., and Jefferys, J. (1998). Electrical coupling underlies high-frequency oscillations in the hippocampus in vitro. *Nature*, 394(6689):189–192.

Şengül, G. and Baysal, U. (2012). An extended kalman filtering approach for the estimation of human head tissue conductivities by using EEG data: a simulation study. *Physiological Measurement*, 33(4):571.

Eulitz, C., Eulitz, H., and Elbert, T. (1997). Differential outcomes from magneto-and electroencephalography for the analysis of human cognition. *Neuroscience letters*, 227(3):185–188.

Fedele, T. (2014). *High-frequency electroencephalography (hf-EEG): non-invasive detection of spike-related brain activity*. Doctoral thesis, Technische Universität Berlin, Fakultät IV - Elektrotechnik und Informatik, Berlin.

Fell, J., Fernandez, G., Klaver, P., Elger, C. E., and Fries, P. (2003). Is synchronized neuronal gamma activity relevant for selective attention? *Brain Research Reviews*, 42(3):265–272.

Fernández-Corazza, M., Beltrachini, L., Von Ellenrieder, N., and Muravchik, C. H. (2013). Analysis of parametric estimation of head tissue conductivities using electrical impedance tomography. *Biomedical Signal Processing and Control*, 8(6):830–837.

Fernández-Corazza, M., Turovets, S., Govyadinov, P., Muravchik, C. H., and Tucker, D. (2016). Effects of head model inaccuracies on regional scalp and skull conductivity estimation using real eit measurements. In Simini, F. and Bertemes-Filho, P., editors, *II Latin American Conference on Bioimpedance*, pages 5–8, Singapore. Springer Singapore.

Fernández-Corazza, M., Turovets, S., Luu, P., Price, N., Muravchik, C. H., and Tucker, D. (2017). Skull modeling effects in conductivity estimates using parametric electrical impedance tomography. *IEEE Transactions on Biomedical Engineering*, 65(8):1785–1797.

Ferree, T., Eriksen, K., and Tucker, D. (2000). Regional head tissue conductivity estimation for improved EEG analysis. *IEEE Trans. Biomed. Eng.*, pages 1584–1592.

Fields, R. D. (2015). A new mechanism of nervous system plasticity: activity-dependent myelination. *Nature Reviews Neuroscience*, 16(12):756–767.

Fisher, R. A. (1936). The use of multiple measurements in taxonomic problems. *Annals of eugenics*, 7(2):179–188.

Fonov, V., Evans, A., Botteron, K., Almli, C., McKinstry, R., Collins, D., and BDCG (2011). Unbiased average age-appropriate atlases for pediatric studies. *NeuroImage*, 54.

Fonov, V., Evans, A., McKinstry, R., Almli, C., and Collins, D. (2009). Unbiased nonlinear average age-appropriate brain templates from birth to adulthood. *NeuroImage*, 47.

Frauscher, B., Bartolomei, F., Kobayashi, K., Cimbalnik, J., van 't Klooster, M. A., Rampp, S., Otsubo, H., Höller, Y., Wu, J. Y., Asano, E., Engel Jr., J., Kahane, P., Jacobs, J., and Gotman, J. (2017). High-frequency oscillations: The state of clinical research. *Epilepsia*, 58(8):1316–1329.

Freeman, W. J. and Zhai, J. (2009). Simulated power spectral density (PSD) of background electrocorticogram (ECoG). *Cognitive neurodynamics*, 3(1):97–103.

Friedman, J. H. (1989). Regularized discriminant analysis. *Journal of the American statistical association*, 84(405):165–175.

Friston, K., Ashburner, J., Kiebel, S., Nichols, T., and Penny, W., editors (2007). *Statistical Parametric Mapping: The Analysis of Functional Brain Images*. Academic Press.

Fuchs, M., Kastner, J., Wagner, M., Hawes, S., and Ebersole, J. (2002). A standardized boundary element method volume conductor model. *Clinical Neurophysiology*, 113(5):702 – 712.

Fuchs, M., Wagner, M., Wischmann, H.-A., Köhler, T., Theißen, A., Drenckhahn, R., and Buchner, H. (1998). Improving source reconstructions by combining bioelectric and biomagnetic data. *Electroencephalography and Clinical Neurophysiology*, 107(2):93 – 111.

Fukunaga, K. (1990). *Introduction to statistical pattern recognition*. Academic Press, Boston, 2nd edition edition.

Gabriel, S., Lau, R., and Gabriel, C. (1996). The dielectric properties of biological tissues: III. parametric models for the dielectric spectrum of tissues. *Physics in Medicine and Biology*, 41(11):2271.

Gao, R., Peterson, E. J., and Voytek, B. (2017). Inferring synaptic excitation/inhibition balance from field potentials. *Neuroimage*, 158:70–78.

Geddes, L. A. and Baker, L. E. (1967). The specific resistance of biological material—a compendium of data for the biomedical engineer and physiologist. *Medical and biological engineering*, 5(3):271–293.

Gevins, A., Smith, M. E., Leong, H., McEvoy, L., Whitfield, S., Du, R., and Rush, G. (1998). Monitoring working memory load during computer-based tasks with EEG pattern recognition methods. *Hum Factors*, 40(1):79–91.

Gilad, O., Horesh, L., and Holder, D. (2007). Design of electrodes and current limits for low frequency electrical impedance tomography of the brain. *Med Bio Eng Comput*, 45:621–633.

Gobbelé, R., Buchner, H., and Curio, G. (1998). High-frequency (600 hz) SEP activities originating in the subcortical and cortical human somatosensory system. *Electroencephalography and Clinical Neurophysiology/Evoked Potentials Section*, 108(2):182 – 189.

Goldenholz, D. M., Ahlfors, S. P., Hämäläinen, M. S., Sharon, D., Ishitobi, M., Vaina, L. M., and Stufflebeam, S. M. (2009). Mapping the signal-to-noise-ratios of cortical sources in magnetoencephalography and electroencephalography. *Human brain mapping*, 30(4):1077–1086.

Goncalves, S., de Munck, J., Heethaar, R., da Silva, F. L., and van Dijk, B. (2000). The application of electrical impedance tomography to reduce systematic errors in the EEG inverse problem—a simulation study. *Physiol. Meas.*, 21:379–393.

Goncalves, S., de Munck, J., Verbunt, J., Bijma, F., Heethaar, R., and da Silva, F. L. (2003). In vivo measurement of the brain and skull resistivities using an EIT-based method and realistic models for the head. *IEEE Trans. Biomed. Eng.*, 50:754–767.

Gramfort, A., Papadopoulo, T., Olivi, E., and Clerc, M. (2010). OpenMEEG: opensource software for quasistatic bioelectromagnetics,. *BioMedical Engineering OnLine*, 45:9.

Güllmar, D., Haueisen, J., and Reichenbach, J. R. (2010). Influence of anisotropic electrical conductivity in white matter tissue on the EEG/MEG forward and inverse solution. a high-resolution whole head simulation study. *Neuroimage*, 51(1):145–163.

Haegens, S., Nácher, V., Luna, R., Romo, R., and Jensen, O. (2011). α-oscillations in the monkey sensorimotor network influence discrimination performance by rhythmical inhibition of neuronal spiking. *Proceedings of the National Academy of Sciences*, 108(48):19377–19382.

Halme, H.-L. and Parkkonen, L. (2016). Comparing features for classification of MEG responses to motor imagery. *PloS one*, 11(12):e0168766.

Häamäläinen, M., Hari, R., Ilmoniemi, R., Knuutila, J., and Lounasmaa, O. (1993). Magnetoencephalography — theory, instrumentation, and applications to noninvasive studies of the working human brain. *Rev. Mod. Phys.*, 65:413–497.

Hansen, S. T., Hauberg, S., and Hansen, L. K. (2016). Data-driven forward model inference for EEG brain imaging. *NeuroImage*, 139:249 – 258.

Hari, R. and Salmelin, R. (1997). Human cortical oscillations: a neuromagnetic view through the skull. *Trends in neurosciences*, 20(1):44–49.

Hasenstaub, A., Shu, Y., Haider, B., Kraushaar, U., Duque, A., and McCormick, D. A. (2005). Inhibitory postsynaptic potentials carry synchronized frequency information in active cortical networks. *Neuron*, 47(3):423–435.

Haueisen, J., Büttner, A., Nowak, H., Brauer, H., and Weiller, C. (1999). The influence of conductivity changes in boundary element compartments on the forward and inverse problem in electroencephalography and magnetoencephalography - der einfluß der änderung der schalenleitfähigkeit bei randelementemodellen auf die vorwärtsrechnung und das inverse problem in elektroenzephalographie und magnetoenzephalographie. *Biomedizinische Technik/Biomedical Engineering*, 44(6):150–157.

Haueisen, J., Ramon, C., Eiselt, M., Brauer, H., and Nowak, H. (1997). Influence of tissue resistivities on neuromagnetic fields and electric potentials studied with a finite element model of the head. *IEEE Transactions on Biomedical Engineering*, 44(8):727–735.

Haufe, S., Dähne, S., and Nikulin, V. V. (2014a). Dimensionality reduction for the analysis of brain oscillations. *NeuroImage*, 101:583–597.

Haufe, S., Meinecke, F., Görgen, K., Dähne, S., Haynes, J.-D., Blankertz, B., and Bießmann, F. (2014b). On the interpretation of weight vectors of linear models in multivariate neuroimaging. *NeuroImage*, 87:96–110.

Haufe, S., Nikulin, V., Ziehe, A., Müller, K., and Nolte, G. (2008). Combining sparsity and rotational invariance in EEG/MEG source reconstruction. *NeuroImage*, 42:726–738.

He, B. J. (2014). Scale-free brain activity: past, present, and future. *Trends in cognitive sciences*, 18(9):480–487.

Henrie, J. A. and Shapley, R. (2005). LFP power spectra in V1 cortex: the graded effect of stimulus contrast. *Journal of neurophysiology*, 94(1):479–490.

Hines, M. L. and Carnevale, N. T. (1997). The neuron simulation environment. *Neural computation*, 9(6):1179–1209.

Holder, D. (2005). *Electric Impedance Tomography*. Institue of Physics Publishing, London.

Holm, A., Lukander, K., Korpela, J., Sallinen, M., and Muller, K. M. (2009). Estimating brain load from the EEG. *ScientificWorldJournal*, 9:639–651.

Holt, G. R., Softky, W. R., Koch, C., and Douglas, R. J. (1996). Comparison of discharge variability in vitro and in vivo in cat visual cortex neurons. *Journal of neurophysiology*, 75(5):1806–1814.

Horesh, L., Bayford, R., Yerworth, R., Tizzard, A., Ahadzi, G., and Holder, D. (2004). Beyond the linear domain - the way forward in MFEIT image reconstruction of the human head. In *proceedings of the ICEBI'04-XII International Conference on Electrical Bio-Impedance joint with EIT-V Electrical Impedance Tomography 2004, ISBN 83-917681-6*, volume 3, pages 683–686.

Hotson, J. and Prince, D. (1980). A calcium-activated hyperpolarization follows repetitive firing in hippocampal neurons. *Journal of Neurophysiology*, 43(2):409–419.

Huang, Y., Dmochowski, J., Su, Y., Datta, A., Rorden, C., and Parra, L. (2013). Automated MRI segmentation for individualized modeling of current flow in the human head. *Journal of Neural Engineering*, 10(6):066004.

Huang, Y., Parra, L., and Haufe, S. (2015). The newyork head—a precise standardized volume conductormodel for EEG source localization and tES targeting. *NeuroImage*, pages –.

Intriligator, J. and Polich, J. (1994). On the relationship between background EEG and the P300 event-related potential. *Biological psychology*, 37(3):207–218.

Jensen, O. and Colgin, L. L. (2007). Cross-frequency coupling between neuronal oscillations. *Trends in cognitive sciences*, 11(7):267–269.

Jirsa, V. K. and Haken, H. (1997). A derivation of a macroscopic field theory of the brain from the quasi-microscopic neural dynamics. *Physica D: Nonlinear Phenomena*, 99(4):503–526.

Jobsis, F. F. (1977). Noninvasive, infrared monitoring of cerebral and myocardial oxygen sufficiency and circulatory parameters. *Science*, 198(4323):1264–1267.

Johnson, J. B. (1925). The Schottky effect in low frequency circuits. *Physical review*, 26(1):71.

Jones, S. R. (2016). When brain rhythms aren't 'rhythmic': implication for their mechanisms and meaning. *Current Opinion in Neurobiology*, 40:72–80.

Kaiser, H. F. (1958). The varimax criterion for analytic rotation in factor analysis. *Psychometrika*, 23(3):187–200.

Kappenman, E. S. and Luck, S. J. (2010). The effects of electrode impedance on data quality and statistical significance in ERP recordings. *Psychophysiology*, 47(5):888–904.

Kirchner, W. K. (1958). Age differences in short-term retention of rapidly changing information. *Journal of Experimental Psychology*, 55:352–358.

Klimesch, W. (2012). Alpha-band oscillations, attention, and controlled access to stored information. *Trends in cognitive sciences*, 16(12):606–617.

Koles, Z. J. (1991). The quantitative extraction and topographic mapping of the abnormal components in the clinical EEG. *Electroencephalogr Clin Neurophysiol*, 79(6):440–447.

Kruskal, J. (1964). Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrika*, 29(1):1–27.

Kybic, J., Clerc, M., Abboud, T., Faugeras, O., Keriven, R., and Papadopoulo, T. (2005). A common formalism for the integral formulations of the forward EEG problem. *IEEE Trans. Med. Img.*, 24:12–28.

Lampl, I. and Yarom, Y. (1993). Subthreshold oscillations of the membrane potential: a functional synchronizing and timing device. *Journal of neurophysiology*, 70(5):2181–2186.

Lawson, C. L. and Hanson, R. J. (1995). *Solving least squares problems*, volume 15. SIAM.

Le Van Quyen, M., Bragin, A., Staba, R., Crépon, B., Wilson, C. L., and Engel, J. (2008). Cell type-specific firing during ripple oscillations in the hippocampal formation of humans. *Journal of Neuroscience*, 28(24):6104–6110.

Ledoit, O. and Wolf, M. (2004). A well-conditioned estimator for large-dimensional covariance matrices. *J Multivar Anal*, 88:365–411.

Lefler, Y., Yarom, Y., and Uusisaari, M. Y. (2014). Cerebellar inhibitory input to the inferior olive decreases electrical coupling and blocks subthreshold oscillations. *Neuron*, 81(6):1389–1400.

Lehtelä, L., Salmelin, R., and Hari, R. (1997). Evidence for reactive magnetic 10-hz rhythm in the human auditory cortex. *Neuroscience letters*, 222(2):111–114.

Lühmann, A. v. (2018). *Multimodal instrumentation and methods for neurotechnology out of the lab*. Doctoral thesis, Technische Universität Berlin, Berlin.

Linkenkaer-Hansen, K., Nikouline, V. V., Palva, J. M., and Ilmoniemi, R. J. (2001). Long-range temporal correlations and scaling behavior in human brain oscillations. *Journal of Neuroscience*, 21(4):1370–1377.

Lotte, F. and Guan, C. (2011). Regularizing common spatial patterns to improve BCI designs: unified theory and new algorithms. *IEEE Trans Biomed Eng*, 58:355–362.

Lowen, S. B. and Teich, M. C. (2005). *Fractal-based point processes*, volume 366. John Wiley & Sons.

Lu, H., Plataniotis, K. N., and Venetsanopoulos, A. N. (2009). Regularized common spatial patterns with generic learning for EEG signal classification. In *2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 6599–6602. IEEE.

Lundström, I. and McQueen, D. (1974). A proposed 1/f noise mechanism in nerve cell membranes. *Journal of theoretical biology*, 45(2):405–409.

Makeig, S., Bell, A. J., Jung, T.-P., and Sejnowski, T. J. (1996). Independent component analysis of electroencephalographic data. In *Advances in neural information processing systems*, pages 145–151.

Malony, A., Salman, A., Turovets, S., Tucker, D., Volkov, V., Li, K., Song, J., Biersdorff, S., Davey, C., Hoge, C., and Hammond, D. (2011). Computational modeling of human head electromagnetics for source localization of milliscale brain dynamics. *Stud Health Technol Inform.*, pages 329–335.

Mantini, D., Perrucci, M. G., Del Gratta, C., Romani, G. L., and Corbetta, M. (2007). Electrophysiological signatures of resting state networks in the human brain. *Proceedings of the National Academy of Sciences*, 104(32):13170–13175.

Marcus, D., Wang, T., Parker, J., Csernansky, J., Morris, J., and Buckner, R. (2007). Open access series of imaging studies (OASIS): Cross-sectional MRI data in young, middle aged, nondemented, and demented older adults. *Journal of Cognitive Neuroscience*, 19(9):1498–1507.

Matsumoto, N., Okamoto, K., Takagi, Y., and Ikegaya, Y. (2016). 3-hz subthreshold oscillations of CA2 neurons in vivo. *Hippocampus*, 26(12):1570–1578.

Mazziotta, J., Toga, A., Evans, A., Fox, P., Lancaster, J., Zilles, K., Woods, R., Paus, T., Simpson, G., Pike, B., Holmes, C., Collins, L., Thompson, P., MacDonald, D., Iacoboni, M., Schormann, T., Amunts, K., Palomero-Gallagher, N., Geyer, S., Parsons, L., Narr, K., Kabani, N., Goualher, G. L., Boomsma, D., Cannon, T., Kawashima, R., and Mazoyer, B. (2001). A probabilistic atlas and reference system for the human brain: International consortium for brain mapping (ICBM). *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 356(1412):1293–1322.

McCormick, D. A. and Pape, H.-C. (1990). Properties of a hyperpolarization-activated cation current and its role in rhythmic oscillation in thalamic relay neurones. *The Journal of physiology*, 431(1):291–318.

Megıas, M., Emri, Z., Freund, T., and Gulyas, A. (2001). Total number and distribution of inhibitory and excitatory synapses on hippocampal CA1 pyramidal cells. *Neuroscience*, 102(3):527–540.

Meijs, J., Weier, O., Peters, M., and van Oosterom, A. (1989). On the numerical accuracy of the boundary element method (EEG application). *Biomedical Engineering, IEEE Transactions on*, 36(10):1038–1049.

Michel, C. M., Murray, M. M., Lantz, G., Gonzalez, S., Spinelli, L., and de Peralta, R. G. (2004). EEG source imaging. *Clinical neurophysiology*, 115(10):2195–2222.

Miklavcic, D., Pavselj, N., and Hart, F. (2006). Electric porperties of tissues. *Wiley Encyclopedia of Biomedical Engineering*.

Miklody, D., Moessmer, P., Dettmann, T., Klinkenberg, K., and Blankertz, B. (2017). Multi-timescale spectra as features for continuous workload estimation in realistic settings. In *Proceedings of The 7th International BCI Conference Graz, 2017*.

Miklody, D., Uitterhoeve, W. M., van Heel, D., Klinkenberg, K., and Blankertz, B. (2016). Maritime cognitive workload assessment. In *International Workshop on Symbiotic Interaction*, pages 102–114. Springer.

Miller, K. J., Honey, C. J., Hermes, D., Rao, R. P., Ojemann, J. G., et al. (2014). Broadband changes in the cortical surface potential track activation of functionally diverse neuronal populations. *Neuroimage*, 85:711–720.

Miller, K. J., Leuthardt, E. C., Schalk, G., Rao, R. P., Anderson, N. R., Moran, D. W., Miller, J. W., and Ojemann, J. G. (2007). Spectral changes in cortical surface potentials during motor movement. *Journal of Neuroscience*, 27(9):2424–2432.

Milton, A. and Pleydell-Pearce, C. W. (2016). The phase of pre-stimulus alpha oscillations influences the visual perception of stimulus timing. *Neuroimage*, 133:53–61.

Mosher, J., Lewis, P., and Leahy, R. (1992). Multiple dipole modeling and localization from spatio-temporal MEG data. *Biomedical Engineering, IEEE Transactions on*, 39(6):541–557.

Munck, J. D., Faes, T., Hermans, A., and Heethaar, R. (1999). A parametric method to resolve the ill-posed nature of the EIT reconstruction problem: A simulation study. *Annals of the New York Academy of Sciences*, 873(1):440–453.

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

Musha, T. and Yamamoto, M. (1997). 1/f fluctuations in biological systems. In *Proceedings of the 19th Annual International Conference of the IEEE Engineering in Medicine and Biology Society.'Magnificent Milestones and Emerging Opportunities in Medical Engineering'(Cat. No. 97CH36136)*, volume 6, pages 2692–2697. IEEE.

Nadasdy, Z. (2010). Binding by asynchrony: the neuronal phase code. *Frontiers in Neuroscience*, 4:51.

Nikulin, V. V. and Brismar, T. (2006). Phase synchronization between alpha and beta oscillations in the human electroencephalogram. *Neuroscience*, 137(2):647–657.

Nikulin, V. V., Nolte, G., and Curio, G. (2011). A novel method for reliable and fast extraction of neuronal EEG/MEG oscillations on the basis of spatio-spectral decomposition. *NeuroImage*, 55(4):1528–1535.

Nissinen, A., Kaipio, J., Vauhkonen, M., and Kolehmainen, V. (2015). Contrast enhancement in eit imaging of the brain. *Physiological measurement*, 37(1):1.

Nissinen, A., Kolehmainen, V., and Kaipio, J. (2011). Compensation of modelling errors due to unknown domain boundary in electrical impedance tomography. *Medical Imaging, IEEE Transactions on*, 30(2):231–242.

Nolte, G. and Dassios, G. (2005). Analytic expansion of the EEG lead field for realistic volume conductors. *Physics in Medicine and Biology*, 50(16):3807.

Novikov, E., Novikov, A., Shannahoff-Khalsa, D., Schwartz, B., and Wright, J. (1997). Scale-similar activity in the brain. *Physical Review E*, 56(3):R2387.

Nunez, P. L. and Srinivasan, R. (2005). *Electric Fields of the Brain: The Neurophysics of EEG*. Oxford University Press, USA, 2 edition.

Nunez, P. L. and Srinivasan, R. (2006). A theoretical basis for standing and traveling brain waves measured with human EEG with implications for an integrated consciousness. *Clinical neurophysiology*, 117(11):2424–2435.

Ogawa, S., Lee, T.-M., Kay, A. R., and Tank, D. W. (1990). Brain magnetic resonance imaging with contrast dependent on blood oxygenation. *proceedings of the National Academy of Sciences*, 87(24):9868–9872.

Ollikainen, J. O., Vauhkonen, M., Karjalainen, P. A., and Kaipio, J. P. (2000). Effects of electrode properties on EEG measurements and a related inverse problem. *Medical engineering & physics*, 22(8):535–545.

Oostendorp, T. F., Delbeke, J., and Stegeman, D. F. (2000). The conductivity of the human skull: results of in vivo and in vitro measurements. *IEEE Transactions on Biomedical Engineering*, 47(11):1487–1492.

Oostenveld, R., Fries, P., Maris, E., and Schoffelen, J. (2011). Fieldtrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data,. *Computational Intelligence and Neuroscience*, 2011:Article ID 156869.

Oostenveld, R. and Praamstra, P. (2001). The five percent electrode system for high-resolution EEG and ERP measurements. *Clinical Neurophysiology*, 112(4):713 – 719.

Ozaki, I. and Hashimoto, I. (2011). Exploring the physiology and function of high-frequency oscillations (HFOs) from the somatosensory cortex. *Clinical neurophysiology*, 122(10):1908–1923.

Pajevic, S., Basser, P. J., and Fields, R. D. (2014). Role of myelin plasticity in oscillations and synchrony of neuronal activity. *Neuroscience*, 276:135–147.

Pearson, K. (1901). LIII. on lines and planes of closest fit to systems of points in space. *Philosophical Magazine Series 6*, 2(11):559–572.

Pedroarena, C. and Llinás, R. (1997). Dendritic calcium conductances generate high-frequency oscillation in thalamocortical neurons. *Proceedings of the National Academy of Sciences*, 94(2):724–728.

Perel, S., Sadtler, P. T., Oby, E. R., Ryu, S. I., Tyler-Kabara, E. C., Batista, A. P., and Chase, S. M. (2015). Single-unit activity, threshold crossings, and local field potentials in motor cortex differentially encode reach kinematics. *Journal of neurophysiology*, 114(3):1500–1512.

Pesaran, B., Pezaris, J. S., Sahani, M., Mitra, P. P., and Andersen, R. A. (2002). Temporal structure in neuronal activity during working memory in macaque parietal cortex. *Nature neuroscience*, 5(8):805–811.

Pfurtscheller, G. and Neuper, C. (2001). Motor imagery and direct brain-computer communication. *Proceedings of the IEEE*, 89(7):1123–1134.

Pfurtscheller, G., Stancak Jr, A., and Edlinger, G. (1997). On the existence of different types of central beta rhythms below 30 hz. *Electroencephalography and clinical neurophysiology*, 102(4):316–325.

Pike, F. G., Goddard, R. S., Suckling, J. M., Ganter, P., Kasthuri, N., and Paulsen, O. (2000). Distinct frequency preferences of different types of rat hippocampal neurones in response to oscillatory input currents. *The Journal of physiology*, 529(1):205–213.

Plöchl, M., Ossandón, J. P., and König, P. (2012). Combining EEG and eye tracking: identification, characterization, and correction of eye movement artifacts in electroencephalographic data. *Frontiers in human neuroscience*, 6:278.

Puil, E., Meiri, H., and Yarom, Y. (1994). Resonant behavior and frequency preferences of thalamic neurons. *Journal of neurophysiology*, 71(2):575–582.

Pykett, I. L. (1982). NMR imaging in medicine. *Scientific American*, 246(5):78–91.

Ramon, C., Schimpf, P., Haueisen, J., Holmes, M., and Ishimaru, A. (2004). Role of soft bone, CSF and gray matter in EEG simulations. *Brain topography*, 16(4):245–248.

Ramoser, H., Müller-Gerking, J., and Pfurtscheller, G. (2000). Optimal spatial filtering of single trial EEG during imagined hand movement. *IEEE Trans Rehabil Eng*, 8(4):441–446.

Rieke, F., Warland, D., Van Steveninck, R. D. R., Bialek, W. S., et al. (1999). *Spikes: exploring the neural code*, volume 7. MIT press Cambridge.

Risner, M. L., Aura, C. J., Black, J. E., and Gawne, T. J. (2009). The visual evoked potential is independent of surface alpha rhythm phase. *Neuroimage*, 45(2):463–469.

Ritter, P. and Becker, R. (2009). Detecting alpha rhythm phase reset by phase sorting: caveats to consider. *Neuroimage*, 47(1):1–4.

Robinson, P. A., Zhao, X., Aquino, K. M., Griffiths, J., Sarkar, S., and Mehta-Pandejee, G. (2016). Eigenmodes of brain activity: Neural field theory predictions and comparison with experiment. *NeuroImage*, 142:79–98.

Russell, G. S., Eriksen, K. J., Poolman, P., Luu, P., and Tucker, D. M. (2005). Geodesic photogrammetry for localizing sensor positions in dense-array EEG. *Clinical Neurophysiology*, 116(5):1130 – 1140.

Sajib, S. Z. K., Jeong, W. C., Kyung, E. J., Kim, H. B., Oh, T. I., Kim, H. J., Kwon, O. I., and Woo, E. J. (2016). Experimental evaluation of electrical conductivity imaging of anisotropic brain tissues using a combination of diffusion tensor imaging and magnetic resonance electrical impedance tomography. *AIP Advances*, 6(6).

Samek, W., Vidaurre, C., Müller, K.-R., and Kawanabe, M. (2012). Stationary common spatial patterns for brain-computer interfacing. *Journal of Neural Engineering*, 9(2):026013.

Sarvas, J. (1987). Basic mathematical and electromagnetic concepts of the biomagnetic inverse problem. *Phys. Med. Biol.*, 32:11–22.

Sauseng, P., Klimesch, W., Stadler, W., Schabus, M., Doppelmayr, M., Hanslmayr, S., Gruber, W. R., and Birbaumer, N. (2005). A shift of visual spatial attention is selectively associated with human EEG alpha activity. *European Journal of Neuroscience*, 22(11):2917–2926.

Schaworonkow, N. and Nikulin, V. V. (2019). Spatial neuronal synchronization and the waveform of oscillations: Implications for EEG and MEG. *PLoS computational biology*, 15(5):e1007055.

Scheer, H. J., Sander, T., and Trahms, L. (2005). The influence of amplifier, interface and biological noise on signal quality in high-resolution EEG recordings. *Physiological measurement*, 27(2):109.

Schirrmeister, R. T., Springenberg, J. T., Fiederer, L. D. J., Glasstetter, M., Eggensperger, K., Tangermann, M., Hutter, F., Burgard, W., and Ball, T. (2017). Deep learning with convolutional neural networks for EEG decoding and visualization. *Human brain mapping*, 38(11):5391–5420.

Schmahmann, J. D., Smith, E. E., Eichler, F. S., and Filley, C. M. (2008). Cerebral white matter. *Annals of the New York Academy of Sciences*, 1142(1):266–309.

Schottky, W. (1926). Small-shot effect and flicker effect. *Physical Review*, 28(1):74.

Shepard, R. N. (1962). The analysis of proximities: multidimensional scaling with an unknown distance function. i. *Psychometrika*, 27(2):125–140.

Sherman, S. M. (2001). Tonic and burst firing: dual modes of thalamocortical relay. *Trends in neurosciences*, 24(2):122–126.

Smith, M. E., Gevins, A., Brown, H., Karnik, A., Du, R., and Gevins, A. S. (2001). Monitoring task loading with multivariate EEG measures during complex forms of human-computer interaction. *Hum Factors*, 43:366–380.

Suffczynski, P., Kalitzin, S., Pfurtscheller, G., and Da Silva, F. L. (2001). Computational model of thalamo-cortical networks: dynamical control of alpha rhythms in relation to focal attention. *International Journal of Psychophysiology*, 43(1):25–40.

Tadel, F., Baillet, S., Mosher, J., Pantazis, D., and Leahy, R. (2011). Brainstorm: A user-friendly application for meg/eeg analysis. *Computational Intelligence and Neuroscience*, 879716.

Tenke, C. E. and Kayser, J. (2005). Reference-free quantification of EEG spectra: combining current source density (CSD) and frequency principal components analysis (fPCA). *Clinical Neurophysiology*, 116(12):2826–2846.

Torgerson, W. (1952). Multidimensional scaling: I. theory and method. *Psychometrika*, 17(4):401–419.

Tuch, D., Wedeen, V., Dale, A., and Belliveau, J. G. J. (2001). Conductivity tensor mapping of the human brain using diffusion tensor MRI. *PNAS*, 98/20:11697–11701.

Turovets, S., Poolman, P., Salman, A., Malony, A., and Tucker, D. (2008). Conductivity analysis for high-resolution EEG. *International Conference on BioMedical Engineering and Informatics (BMEI)*, pages 386 – 393.

Valdés-Hernández, P., von Ellenrieder, N., Ojeda-Gonzalez, A., Kochen, S., Alemán-Gómez, Y., Muravchik, C., and Valdés-Sosa, P. (2009). Approximate average head models for EEG source imaging. *Journal of Neuroscience Methods*, 185(1):125 – 132.

van Dijk, H., van der Werf, J., Mazaheri, A., Medendorp, W. P., and Jensen, O. (2010). Modulations in oscillatory activity with amplitude asymmetry can produce cognitively relevant event-related responses. *Proceedings of the National Academy of Sciences*, 107(2):900–905.

Van Hoey, G., Vanrumste, B., D'have, M., Van de Walle, R., Lemahieu, I., and Boon, P. (2000). Influence of measurement noise and electrode mislocalisation on eeg dipole-source localisation. *Medical and Biological Engineering and Computing*, 38(3):287–296.

Vidaurre, C., Kawanabe, M., von Bünau, P., Blankertz, B., and Müller, K.-R. (2011). Toward unsupervised adaptation of LDA for brain-computer interfaces. *IEEE Trans Biomed Eng*, 58(3):587 –597.

von Bünau, P., Meinecke, F. C., Király, F., and Müller, K.-R. (2009). Finding stationary subspaces in multivariate time series. *Physical Review Letters*, 103:214101.

Vorwerk, J., Cho, J.-H., Rampp, S., Hamer, H., Knösche, T. R., and Wolters, C. H. (2014). A guideline for head volume conductor modeling in EEG and MEG. *NeuroImage*, 100(0):590 – 607.

Vorwerk, J., Clerc, M., Burger, M., and Wolters, C. (2012). Comparison of boundary element and finite element approaches to the EEG forward problem. *Biomed Tech*, pages 795 – 798.

Wagner, S., Lucka, F., Vorwerk, J., Herrmann, C. S., Nolte, G., Burger, M., and Wolters, C. H. (2016). Using reciprocity for relating the simulation of transcranial current stimulation to the EEG forward problem. *Neuroimage*, 140:163–173.

Wang, B., Ke, W., Guang, J., Chen, G., Yin, L., Deng, S., He, Q., Liu, Y., He, T., Zheng, R., et al. (2016). Firing frequency maxima of fast-spiking neurons in human, monkey, and mouse neocortex. *Frontiers in cellular neuroscience*, 10:239.

Wang, X.-J. and Buzsáki, G. (1996). Gamma oscillation by synaptic inhibition in a hippocampal interneuronal network model. *Journal of neuroscience*, 16(20):6402–6413.

Wang, Y. and Gotman, J. (2001). The influence of electrode location errors on EEG dipole source localization with a realistic head model. *Clinical Neurophysiology*, 112(9):1777–1780.

Welch, P. (1967). The use of fast fourier transform for the estimation of power spectra: a method based on time averaging over short, modified periodograms. *IEEE Transactions on audio and electroacoustics*, 15(2):70–73.

Winkler, I., Haufe, S., and Tangermann, M. (2011). Automatic classification of artifactual ICA-components for artifact removal in EEG signals. *Behavioral and Brain Functions*, 7(1):1–15.

Wolpaw, J. R. and McFarland, D. J. (2004). Control of a two-dimensional movement signal by a noninvasive brain-computer interface in humans. *Proceedings of the national academy of sciences*, 101(51):17849–17854.

Wolters, C., Anwander, A., Tricoche, X., Weinstein, D., Koch, M., and 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.

Wróbel, A. (2000). Beta activity: a carrier for visual attention. *Acta neurobiologiae experimentalis*, 60(2):247–260.

Yger, F., Berar, M., and Lotte, F. (2016). Riemannian approaches in brain-computer interfaces: a review. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 25(10):1753–1762.

Yordanova, J., Kolev, V., and Polich, J. (2001). P300 and alpha event-related desynchronization (erd). *Psychophysiology*, 38(1):143–152.

Yvert, B., Bertrand, O., Thévenet, M., Echallier, J., and Pernier, J. (1997). A systematic evaluation of the spherical model accuracy in EEG dipole localization. *Electroencephalography and Clinical Neurophysiology*, 102(5):452 – 459.

Zayed, A. I. (2019). *Handbook of function and generalized function transformations*. CRC press.

Zhan, X., Cox, C. L., Rinzel, J., and Sherman, S. M. (1999). Current clamp and modeling studies of low-threshold calcium spikes in cells of the cat's lateral geniculate nucleus. *Journal of neurophysiology*, 81(5):2360–2373.

Zhang, Y. and Haniza, A. (2006). Vertex-based anisotropic smoothing of 3D mesh data. In *Electrical and Computer Engineering, 2006. CCECE '06. Canadian Conference on*, pages 202–205.

List of Figures

List of Tables