

Dissertation

**New mathematics for EEG/MEG source analysis and
individually optimized transcranial electric stimulation and
application in a case study of focal epilepsy**

Malte Bernhard Höltershinken

2025

New mathematics for EEG/MEG source analysis and individually optimized transcranial electric stimulation and application in a case study of focal epilepsy

Inauguraldissertation zur Erlangung des Doktorgrades der Naturwissenschaften
- Dr. rer. nat. -
im Fachbereich Mathematik und Informatik der Mathematisch-Naturwissenschaftlichen
Fakultät der Universität Münster als interdisziplinäre Promotion in Zusammenarbeit mit der
Medizinischen Fakultät

vorgelegt von
Malte Bernhard Höltershinken
aus Lingen

Münster, 2025

Dekan:

Prof. Dr. Arthur Bartels

Erstgutachter:

Prof. Dr. Christian Engwer

Zweitgutachter:

Prof. Dr. Carsten H. Wolters

Drittgutachter:

Prof. Dr. Martin Burger

Tag der mündlichen Prüfung:

Tag der Promotion:

Abstract

Neural activity generates electric currents. These currents result in electric and magnetic fields that can be measured outside of the head. Using these non-invasive measurements to estimate active brain regions is a fundamental building block in many medical and neuroscientific applications. Furthermore, one can also stimulate the brain by applying electromagnetic fields to modulate neural activity, which has been shown to have the potential of inducing a positive effect in a number of medical conditions.

Both the estimation of active brain areas and the prediction of brain stimulation effects rely on an accurate simulation of electromagnetic fields inside and outside of the head. The majority of this thesis is concerned with numerical strategies to perform this simulation. Concretely, we present an improvement on the subtraction approach for finite element method (FEM) based simulation of electromagnetic fields. While the subtraction approach is mathematically rigorous and yields accurate simulations, its high computational demand has prevented its practical use. In this thesis, we present the so-called local subtraction approach, which is as rigorous as the subtraction approach, but vastly more efficient. We show that the local subtraction approach can outperform the accuracy of existing state-of-the-art approaches and demonstrate the viability of the approach in realistic settings.

Additionally, we present a new derivation of the reciprocity laws of quasi-static bioelectromagnetism. While existing derivations are based on complex computations, we will introduce a functional analytic framework that yields a unified view on the different expressions of reciprocity laws and enables a straightforward and conceptually clear derivation.

Furthermore, we show that two of the most popular approaches for estimating active brain areas, namely dipole scanning and sLORETA, are equivalent. In single-source scenarios, we will moreover show that dipole scanning is equivalent to SAM and NAI beamforming.

Finally, we present the electrical brain stimulation of an epilepsy patient, where the methods mentioned above were applied with the goal of reducing the patient's epileptic activity. This stimulation did not result in a significant reduction of epileptic spikes. The absence of a positive stimulation effect, however, informed the further treatment process by influencing the decision on whether to proceed with the implantation of a stimulation device.

Zusammenfassung

Neuronale Aktivität erzeugt elektrische Ströme. Diese Ströme resultieren in elektrischen und magnetischen Feldern, die außerhalb des Kopfes gemessen werden können. Diese nichtinvasiven Messungen können nun benutzt werden, um die zugrundeliegenden aktiven Hirnregionen zu schätzen, was eine wichtige Komponente in vielen medizinischen und neurowissenschaftlichen Anwendungen ist. Darüber hinaus kann man das Hirn auch stimulieren, indem von außen elektromagnetische Felder angelegt werden. Dies kann neuronale Aktivität modulieren, was einen positiven Effekt auf verschiedene Erkrankungen haben kann.

Sowohl die Rekonstruktion neuronaler Aktivität als auch die Vorhersage von Effekten einer Hirnstimulation hängen von einer genauen Simulation der elektromagnetischen Felder innerhalb und außerhalb des Kopfes ab. Der Großteil dieser Arbeit beschäftigt sich mit numerischen Verfahren für die Durchführung dieser Simulation. Konkret stellen wir eine Verbesserung des Subtraktionsansatzes für die Finite-Element-Methode (FEM) basierte Simulation von elektromagnetischen Feldern vor. Der Subtraktionsansatz ist mathematisch rigoros und produziert genaue Simulationen, hat jedoch das Problem, durch einen zu hohen Berechnungsaufwand nicht praxistauglich zu sein. In dieser Arbeit stellen wir den sogenannten lokalen Subtraktionsansatz vor, der genauso rigoros wie der Subtraktionsansatz ist, jedoch eine wesentlich effizientere Berechnung ermöglicht. Wir zeigen, dass der lokale Subtraktionsansatz die etablierten Methoden in Bezug auf Genauigkeit übertreffen kann, und darüber hinaus in realistischen Szenarien einsetzbar ist.

Darüber hinaus präsentieren wir eine neue Herleitung der Reziprozitätsgesetze des quasistatischen Bioelektromagnetismus. Während die bekannten Herleitungen auf umständlichen Rechnungen basieren, stellen wir eine einfache funktionalanalytische Perspektive vor, die die verschiedenen Formen von Reziprozität vereinheitlicht und eine unkomplizierte und konzeptionell klare Herleitung ermöglicht.

Außerdem zeigen wir, dass zwei der bekanntesten Ansätze zur Rekonstruktion neuronaler Aktivität, dipole scanning und sLORETA, äquivalent sind. Zusätzlich zeigen wir das dipole scanning in Ein-Quell-Szenarios ebenfalls mit SAM und NAI beamforming Ansätzen identifiziert werden kann.

Abschließend stellen wir die elektrische Hirnstimulation eines Epilepsie-Patienten vor. In diesem Experiment wurden die oben genannten Methoden angewendet mit dem Ziel die epileptische Aktivität des Patienten zu reduzieren. Im Rahmen dieser Stimulation konnten wir keine Reduktion der epileptischen Aktivität erreichen. Das Ausbleiben eines positiven Stimulationseffektes hat allerdings Informationen für den weiteren Behandlungsprozess geliefert, indem die Entscheidung über eine potenzielle Implantation eines Stimulationsgerätes beeinflusst wurde.

Acknowledgments

This thesis did not emerge from the void, but was only possible due to the support, guidance, and trust of numerous people and institutions. I want to sincerely thank everyone who aided me during the past four years. I, in particular, want to express my gratitude to the following people.

- Carsten H. Wolters, for expanding my horizon past the boundary of theoretical mathematics, for giving me both the opportunity to pursue this interdisciplinary work and the freedom to choose topics I believe in, and for always being available for guidance and discussion.
- Christian Engwer, for guiding me through the world of numerical mathematics, for teaching me how to translate mathematical ideas into usable implementations, and for actually taking the time to go through code line by line to help me overcome the problems I was facing, without which a large portion of this thesis would not have been possible.
- Martin Burger, for reviewing this thesis.
- Tim Erdbrügger, for always being open to questions and discussions, and for being a wonderful travelling companion during all of the conferences we visited.
- Fatih Karpuz, for always having an open ear for my mathematical and computer woes, for offering thoughtful feedback, and for proofreading large parts of this thesis.
- Nils-Arne Dreier, who, when I first started my master's thesis, was always available for mathematical and technical questions, and showed great patience in helping me get accustomed to programming in DUNE and DUNEuro.
- Kanjana Unnwongse, Tim Wehner, Stefan Rampp, and Cindy Gebhardt for enabling the tDCS stimulation experiment we present in this thesis, and for finding the time in their busy schedules to aid us in its execution and evaluation.
- The PerEpi project partners, in particular, Sampsa Pursiainen, Fabrice Wallois, Chenming Yao, Alena Buyx, and Anna Sierawska.
- The DFG project partners, in particular, Jan-Ole Radecke, Yvonne Buschermöhle, Rebecca Lencer, and Joachim Gross.
- Everyone at the IBB for creating an environment that made me look forward to coming to work each day.
- Saskia, for persevering through countless rubberducking sessions, for providing a much-needed counterbalance in the stressful phases of the PhD, and for proofreading this thesis.

-
- My parents, Heinz and Annette, for enabling me to pursue my love of mathematics, for supporting me in every way imaginable, and for giving me the security of knowing that there will always be someone who has my back.
 - My grandfather, Bernhard Höltershinken, for having a significant part in raising me to be the person I am today.

My PhD studies were financed by different funding agencies. I also want to express my gratitude to them.

- ERA PerMed as project ERAPERMED2020-227, PerEpi, where the local funding agency was the Bundesministerium für Gesundheit (BMG), via project ZMI1-2521FSB006
- Deutsche Forschungsgemeinschaft (DFG), projects WO1425/10-2 and WO1425/11-1
- Deutscher Akademischer Austauschdienst (DAAD), via DAAD project 57756407

Contents

1	Introduction	8
2	The forward problems of bioelectromagnetism	11
2.1	The Measurement Forward Problems	12
2.2	The Stimulation Forward Problems	15
2.3	Choosing an approach for solving the forward problems	22
2.4	Deriving FEM formulations for the stimulation forward problems	24
2.4.1	The tDCS forward problem	24
2.4.2	The TMS forward problem	31
2.5	Deriving FEM formulations for the measurement forward problems	33
2.5.1	Direct approaches for the EEG forward problem	34
2.5.2	Subtraction approaches for the EEG forward problem	50
2.5.3	The MEG forward problem	65
2.6	Computing the local subtraction right-hand side	71
2.6.1	Introduction	71
2.6.2	Deriving the infinity potential	73
2.6.3	Right-hand side analytical expressions in the isotropic case	78
2.6.4	Right-hand side analytical expressions in the anisotropic case	99
2.7	Further Implementation Aspects	104
2.7.1	Constructing the local subtraction approach	105
2.7.2	Integration orders for MEG postprocessing	106
2.7.3	Source space construction	110
2.7.4	Setting up a continuous integration pipeline	117
2.8	Numerical investigations	122
2.8.1	Volume conductor models used for validation	123
2.8.2	The local subtraction patch construction	126
2.8.3	Comparing direct approaches for the forward problems	129
2.8.4	Comparing direct and subtraction-type approaches	132
2.8.5	Local subtraction approaches in realistic head models	137
2.9	The reciprocity laws of quasi-static bioelectromagnetism	140
2.9.1	The general reciprocity principle	140
2.9.2	The reciprocity of EEG and tES	142
2.9.3	The reciprocity of MEG and TMS	148
2.9.4	Relation of reciprocity to transfer matrix approaches	153

3	The EEG and MEG inverse problems	156
3.1	Notations and Definitions	159
3.1.1	Lead fields	159
3.1.2	Dipole scanning	160
3.1.3	(Generalized) sLORETA	160
3.1.4	Beamforming	162
3.2	Dipole scanning and sLORETA	167
3.2.1	Identification of sLORETA and dipole scanning	167
3.2.2	The impact of noise	169
3.3	Beamforming investigations	177
3.3.1	Beamforming and dipole scanning	177
3.3.2	Reconstructing source orientations with beamforming	187
4	Source reconstruction and tDCS stimulation of an epilepsy patient	197
4.1	Stimulation montage optimization	200
4.2	Source reconstruction of an epilepsy patient	209
4.2.1	Constructing a volume conductor model	210
4.2.2	Estimating the generators of interictal epileptic activity	219
4.3	TDCS stimulation of an epilepsy patient	224
4.3.1	Description of the stimulation montages	224
4.3.2	Results of the stimulation	226
4.4	Discussion and patient outlook	231
4.5	Conclusion	234

Chapter 1

Introduction

This thesis is concerned with the modeling and interpretation of electromagnetic fields in the context of the human brain. This theme is specialized into multiple subfields.

Firstly, neural activity generates electric currents, which in turn result in electric and magnetic fields that can be measured non-invasively outside of the head [83]. The measuring of electric potentials due to neural activity is called *electroencephalography*, or *EEG*, and measuring the corresponding magnetic field is called *magnetoencephalography*, or *MEG*. One of the central benefits of EEG and MEG compared to other neuroimaging modalities, such as e.g., fMRI, is their high temporal resolution, enabling the sampling of neural activity in the millisecond range [80]. This allows for a multitude of noteworthy applications, both in basic neuroscience, such as, e.g., [184], as well as in clinical practice, such as, e.g., [157]. Many of these applications require accurate knowledge about the relationship between the neural activity and the electromagnetic fields it generates. This is, in particular, the case in the so-called *EEG and MEG inverse problem*, where one tries to estimate the underlying neural activity from the non-invasively measured electric and magnetic fields. The process of performing this estimation is, depending on the author, called either *source imaging/reconstruction/estimation*, and, as the inverse problem does not possess a unique solution, it is *ill-posed* [83]. Simulating the electric and magnetic fields generated by a given neural activity, on the other hand, does have a unique solution, and accurate approaches for this so-called *EEG/MEG forward problem* are an essential prerequisite for a reliable source estimation [216]. Both the EEG/MEG forward and inverse problems are areas of active research, each with a plethora of subtopics that are being investigated. Recent examples include [91], [114], or [220], which is, however, by no means an exhaustive list.

While in EEG and MEG, one passively measures the electromagnetic fields generated by neural activity, one can also go the other way around and try to actively influence neural activity by externally applying electromagnetic fields. The counterpart of EEG is called *transcranial electrical stimulation*, or *tES*, and consists of injecting and extracting electric current into/from a subject's head, typically via electrodes placed on the subject's scalp [165]. It has been shown that tES can have a beneficial effect on different medical conditions, such as depression [24] or focal epilepsy [104, 5]. Furthermore, it has been demonstrated that a mathematical optimization of the tES current injection, that is adapted to the subject's individual head anatomy and based on an accurate simulation of the electric field a current injection would generate, can, in comparison with standard neurophysiological procedures, yield an improved stimulation efficacy [105]. The simulation of the electric field a given current injection would generate is typically called the *tES*

forward problem, and the process of finding an “optimal” current injection is typically called *tES montage optimization*. Especially the montage optimization is an area of active research, with a number of relatively recent publications, such as, e.g., [69, 106, 60]. On the magnetic side, the counterpart to MEG is called *transcranial magnetic stimulation*, or *TMS*. In TMS, a coil is positioned close to the subject’s head and supplied with a time-varying current, which produces an accompanying time-varying electric field in the vicinity of the coil, which can then have an effect on the part of the brain below the coil [202]. Similar to tES, TMS has been demonstrated to have a positive effect on certain medical conditions, such as post-stroke dysphagia [113] or depression [34]. Also similar to tES, an accurate prediction and optimization of TMS stimulation effects requires the faithful modeling of the induced electric field, which depends both on the individual head anatomy as well as on the coil geometry [47, 202]. The problem of modeling the electric field induced by TMS stimulation is called the *TMS forward problem*.

This sets the stage for the content of this thesis. All of the topics mentioned above are touched upon by the following investigations.

The bulk of this thesis consists of investigating finite element approaches for the different forward problems introduced above, which is the content of chapter 2. Although we discuss a number of subtopics about forward modeling, the heart of this chapter is the derivation, theoretical investigation, and numerical evaluation of the *local subtraction approach* for EEG and MEG forward modeling, which was the content of our paper [91]. This approach is a generalization of and improvement on the classical subtraction approach from, e.g., [18] and [239], and, in comparison with the classical subtraction approach, possesses a greatly improved accuracy and computational efficiency. This thesis, in particular, significantly extends on the presentation given in [91], as here we also extend the local subtraction approach to be applicable in the context of the *complete electrode model*, and we additionally give a detailed derivation of the analytical expressions used in the assembly of the finite element right-hand side. Another topic we want to particularly highlight is our new framework for the derivation and interpretation of the so-called *reciprocity laws of bioelectromagnetism*, which we present in section 2.9. While the results presented in this section are already known in the bioelectromagnetism community [205, 140], we present a new derivation, which significantly decreases the complexity of the argument and easily generalizes to more complicated settings. Finally, we also want to highlight that in section 2.5.1 we present an in-depth derivation and mathematical discussion of established approaches for the EEG and MEG forward problems, in particular of the Venant approach [218, 28, 143] and the Whitney type approaches [162, 15], which showcases a number of results that are, to the best of our knowledge, not known in the bioelectromagnetism community.

In chapter 3, we then discuss aspects of the EEG and MEG inverse problems. In this chapter, we will not introduce a new approach, but instead focus on investigating mathematical properties of established approaches. Concretely, we will discuss dipole scans [68], (generalized) sLORETA [149, 150], and beamforming approaches [190]. The chapter then, in essence, consists of two subtopics. First, we will investigate the relations between these approaches. Concretely, we will show that dipole scans are exactly the same as generalized sLORETA approaches. Furthermore, we will show that in scenarios where the neural activity can be adequately modeled as originating from a single dipolar source [40], dipole scans and beamforming will yield equivalent source estimations. These investigations are an expanded and more mature version of the preprint [95]. Secondly, we present investigations into whether so-called *scalar beamformers* are able to reconstruct the orientation of neural sources. This is, in essence, the content of our paper [30]. We, however, want to emphasize that, based on the aforementioned identification of dipole scanning and beamforming, we completely revised and extended the proofs from [30], resulting

in a greater conceptual clarity and simpler arguments.

In chapter 4, we then focus on the application of electrical brain stimulation in focal epilepsy. In section 4.1, we start off on the theoretical side by discussing different established approaches for montage optimization [106, 60], investigating their mathematical properties, and finally suggesting a modification of the algorithm from [106] for use in practical applications. In the remainder of the chapter, we then bring together forward modeling, inverse modeling, and tES montage optimization for the brain stimulation of a focal epilepsy patient. This application is part of the PerEpi project,¹ whose goal was the investigation of the potential of tES stimulation in the treatment, but importantly also in the diagnostic process, of focal epilepsy. The experiment presented in this thesis is thus embedded in the PerEpi project, and continues the study started in [5] and [55]. In section 4.2, we will introduce the patient and show how the patient’s epileptic activity was used to estimate the so-called *irritative zone* of the patient. Furthermore, we will present a detailed description of our data processing pipeline [54], and in particular of the modifications to this pipeline we introduced in this thesis. In sections 4.3 and 4.4, we will then describe the brain stimulation experiment, present its outcome, and discuss the experimental result and future outlook of the patient.

¹<https://www.medizin.uni-muenster.de/perepi/startseite.html>

Chapter 2

The forward problems of bioelectromagnetism

In this chapter, we will introduce our new developments regarding the EEG and MEG forward problems. The central contribution we want to present is a new strategy for incorporating point dipoles into a finite element approach, called the *local subtraction approach*. This result was published in [91] for the so-called point electrode model. In this thesis, we will present a more in-depth discussion of this approach, as well as its extension to the so-called complete electrode model. Furthermore, we will discuss a number of miscellaneous topics regarding the practice of solving the forward problem, such as the construction of a source space.

Additionally, we will present a new derivation of the so-called *reciprocity principles of bioelectromagnetism*. These principles establish relationships between measurement modalities, such as EEG and MEG, and stimulation modalities, such as tES and TMS. Loosely speaking, these principles enable an exchange of the roles of “sensor” and “source”. Previously, these relationships were derived using complicated computations. We will show how they can be derived as direct corollaries from a general, strikingly simple, functional analytic principle. To properly discuss these topics, we will also introduce the mathematical description of the stimulation forward problems in this chapter.

The starting points of all the following derivations are Maxwell’s equations, which describe the electric and magnetic fields by the set of partial differential equations

$$\operatorname{div}(E) = \frac{\rho}{\epsilon_0} \quad (2.1)$$

$$\operatorname{rot}(E) = -\frac{\partial B}{\partial t} \quad (2.2)$$

$$\operatorname{div}(B) = 0 \quad (2.3)$$

$$\operatorname{rot}(B) = \mu_0 \cdot \left(j + \epsilon_0 \frac{\partial E}{\partial t} \right) \quad (2.4)$$

where E denotes the electric field, ρ the charge density, ϵ_0 the vacuum permittivity, B the magnetic flux density, μ_0 the vacuum permeability, and j the current density. We refer to [141] for an introduction to these equations and their role in electromagnetism.

2.1 The Measurement Forward Problems

In the EEG and MEG forward problems, our goal is to simulate the electric and magnetic sensor measurements generated by some given neural activity. Here, we want to derive the corresponding strong problem formulations.

The first step consists of simplifying equations (2.1) - (2.4). Depending on the application, there are different approaches to perform this simplification. Later on, when discussing transcranial magnetic stimulation (TMS), we will see that $-\frac{\partial B}{\partial t}$ in (2.2) cannot be neglected in that setting. In the context of EEG and MEG forward modeling, however, the common strategy in bioelectromagnetism is to ignore the time derivatives in (2.2) and (2.4), which the community refers to as the *quasistatic approximation*. For a justification of this approach, we refer to [83]. Following [83], using this approximation on (2.2) first implies that $\text{rot}(E) = 0$, and we can hence write $E = -\nabla u$ for a scalar potential u , called the *electric potential*. Then, we can derive from (2.4) that

$$0 = \text{div}(\text{rot}(B)) = \mu_0 \cdot \text{div}(j) \iff \text{div}(j) = 0.$$

Now, the current density j inside the head is modeled as originating from the neural activity, called j^P , and the ohmic so-called *volume currents* induced by the electric field, i.e. σE , where σ denotes the electrical conductivity. More concretely, one models the current density as $j = j^P + \sigma E$, which together with the computation above implies

$$0 = \text{div}(j) = \text{div}(j^P) - \text{div}(\sigma \nabla u) \iff \text{div}(\sigma \nabla u) = \text{div}(j^P).$$

Let Ω denote the head domain, and η its outer normal. One then models the conductivity of air as 0. Furthermore, one assumes that the neural activity is confined to the brain, which implies that $j^P = 0$ on a sufficiently small environment of $\partial\Omega$.

The electric potential u is now measured using electrodes. One can employ different approaches to model these measurements. The most commonly used approach is the so-called *point electrode model*, or *PEM*, where one assumes that the presence of the electrodes does not affect the electric potential u , but simply measures this potential at a set of predefined points. More concretely, given that the air is assumed to have conductivity 0, and the neural activity is constrained to the brain, we have in this model that on $\partial\Omega$

$$0 = \langle j, \eta \rangle = -\langle \sigma \nabla u, \eta \rangle \iff \langle \sigma \nabla u, \eta \rangle = 0.$$

Definition 2.1.1 (EEG forward problem, PEM model). Let Ω denote the head domain with outer normal η . Let σ denote the electrical conductivity. Then, the electric potential $u : \Omega \rightarrow \mathbb{R}$ due to the neural activity j^P is described by

$$\begin{aligned} \text{div}(\sigma \nabla u) &= \text{div}(j^P) && \text{in } \Omega \\ \langle \sigma \nabla u, \eta \rangle &= 0 && \text{on } \partial\Omega \end{aligned}$$

A more realistic approach is the so-called *complete electrode model*, or *CEM*, which was originally introduced in the context of electrical impedance tomography [196]. This approach is based on the observation that attaching an electrode to the head surface adds additional possible paths for the electric current, which can now also flow *through* the electrodes. These currents are called *shunting currents*, and we see that the act of performing an EEG measurement has an effect on the measured potentials by distorting the current flow in the head.

Assume that we have $L \in \mathbb{N}$ electrodes. Let $e_1, \dots, e_L \subset \partial\Omega$ denote their contact interfaces with the head domain. We assume that the contact interfaces are pairwise disjoint. In [196], the authors then propose to model each electrode as a perfect conductor, which implies that the potential inside the electrode is constant. Furthermore, using physical reasoning, they propose that on each contact interface e_l , $1 \leq l \leq L$, the current flowing from the head into the electrode is proportional to the potential drop between the electrode and the head. More concretely, if u denotes the potential inside the head and U_l the potential in the l -th electrode, there exists $z_l \in \mathbb{R}$ such that

$$U_l - u = z_l \cdot \langle \sigma \nabla u, \eta \rangle \quad \text{on } e_l.$$

This proportionality factor z_l is then called *effective contact impedance*, or *ECI*, of the l -th electrode. Note that this naming convention is slightly misleading, as e.g. the SI units of z_l are $\Omega \text{ m}^2$, and not Ω , as one might expect from a quantity called impedance.

Furthermore, as no current is inserted or extracted from the head, the net current over each electrode must be zero, i.e., for $1 \leq l \leq L$ we have $\int_{e_l} \langle \sigma \nabla u, \eta \rangle dS = 0$. On the remaining head, one still assumes a zero-outflow condition, i.e. $\langle \sigma \nabla u, \eta \rangle = 0$.

Definition 2.1.2 (EEG forward problem, CEM model). Let Ω denote the head domain with outer normal η . Let σ denote the electrical conductivity. Assume that we employ L electrodes with pairwise disjoint contact interfaces $e_1, \dots, e_L \subset \partial\Omega$ and effective contact impedances $z_1, \dots, z_L \in \mathbb{R}$. Then, the electric potential $u : \Omega \rightarrow \mathbb{R}$ in the head and the potentials $U_1, \dots, U_L \in \mathbb{R}$ in the electrodes due to the neural activity j^P are described by

$$\begin{aligned} \operatorname{div}(\sigma \nabla u) &= \operatorname{div}(j^P) && \text{in } \Omega \\ \langle \sigma \nabla u, \eta \rangle &= 0 && \text{on } \partial\Omega \setminus \bigcup_{l=1}^L e_l \\ z_l \cdot \langle \sigma \nabla u, \eta \rangle &= U_l - u && \text{on } e_l \text{ for } 1 \leq l \leq L \\ \int_{e_l} \langle \sigma \nabla u, \eta \rangle dS &= 0 && \text{for } 1 \leq l \leq L \end{aligned}$$

In [196], it is shown that out of all electrode models the authors compare, which include the PEM model, the CEM model is closest to the real measurement data.

In EEG and MEG forward modeling, one almost always employs the point electrode model. The reason for this is that one expects that the CEM only has a negligible influence on the simulation results, as was, e.g., shown in the comparison study [160]. Note, however, that in this study, the authors investigate electrodes placed on the scalp. In cases where the distance of the electrodes to the neural sources is much smaller, such as in StereoEEG, where electrodes are inserted directly into the patient's brain [142], the realistic modeling of the electrode geometry and the behaviour of the electric potential at the electrode interface will have a larger impact on the simulation accuracy. We thus argue that it is important to also have the CEM model available, as it enables the treatment of more complicated cases than the PEM. Additionally, even in cases where one suspects that the PEM model might be sufficient, it is beneficial to have the CEM model available, as this enables a verification of the hypothesis that the PEM is sufficient by means of comparing the results of the two.

Up until now, we have derived two different approaches for describing the electric potential. Additionally, we also want to simulate the magnetic field generated by neural activity. To derive

a corresponding approach, we first want to remind the reader that for a differentiable scalar field φ and a differentiable vector field F on \mathbb{R}^3 , we have

$$\operatorname{rot}(\varphi \cdot F) = \varphi \cdot \operatorname{rot}(F) + \nabla \varphi \times F \quad (2.5)$$

$$\operatorname{rot}(\operatorname{rot}(F)) = \nabla (\operatorname{div}(F)) - \Delta F, \quad (2.6)$$

as can be verified by direct computation. Furthermore, we want to remind the reader that if we define $N(z) = -\frac{1}{4\pi} \cdot \frac{1}{\|z\|}$ and $\rho : \mathbb{R}^3 \rightarrow \mathbb{R}$ is regular enough, we have

$$\Delta \left(x \mapsto \int_{\mathbb{R}^3} N(x-y) \rho(y) dV(y) \right) = \rho,$$

as is e.g. shown in [64], §16, Satz 1. We can now compute for the magnetic flux density

$$\Delta B = \nabla (\operatorname{div}(B)) - \operatorname{rot}(\operatorname{rot}(B)) \stackrel{(2.3)}{=} -\operatorname{rot}(\operatorname{rot}(B)) \stackrel{\text{quasistatic (2.4)}}{=} -\mu_0 \cdot \operatorname{rot}(j). \quad (2.7)$$

We assume that j is compactly supported. Furthermore, based on physical intuition, we assume that $B(x) \xrightarrow{x \rightarrow \infty} 0$. Then, B is uniquely determined by (2.7), see e.g. [64], §16, Corollar 2, and we have for $x \in \mathbb{R}^3$ that

$$\begin{aligned} B(x) &= \int_{\mathbb{R}^3} N(x-y) \cdot (-\mu_0 \cdot \operatorname{rot}(j)(y)) dV(y) \\ &= \frac{\mu_0}{4\pi} \cdot \int_{\mathbb{R}^3} \frac{1}{\|x-y\|} \cdot \operatorname{rot}(j)(y) dV(y) \\ &= \frac{\mu_0}{4\pi} \cdot \int_{\mathbb{R}^3} - \left(\nabla_y \frac{1}{\|x-y\|} \right) \times j dV(y) \\ &= \frac{\mu_0}{4\pi} \cdot \int_{\mathbb{R}^3} j \times \frac{x-y}{\|x-y\|^3} dV(y). \end{aligned}$$

This last equation is commonly called the *Biot-Savart law*, and it is typically taken as the starting point for solving the MEG forward problem.

Now, the current j has contributions from the neural activity j^P , and from the ohmic volume currents that arise due to the electric field generated by j^P . Inside the head domain Ω , these last currents are given by $-\sigma \nabla u$. In a PEM approach, these are already all the relevant currents. For a CEM approach, there are, strictly speaking, also contributions from the shunting currents inside the electrodes. It is, however, well-known in the MEG community that, due to the quick decay of the volume current strength and the insulating effect of the skull, the contributions of the volume currents in the skull and the scalp to the neural magnetic field are negligible [6, 91, 83]. But then the shunting currents between the scalp and the electrodes are also negligible, and we will, hence, in this thesis, ignore them in the context of the computation of the magnetic field. We are, in fact, not aware of a single publication where shunting currents inside the electrodes were taken into account for the simulation of neural magnetic fields.

Definition 2.1.3 (MEG forward problem). Let Ω denote the head domain. Let σ denote the conductivity, j^P the neural activity, and $u : \Omega \rightarrow \mathbb{R}$ the electric potential in the head generated by j^P . Then the MEG forward problem consists of computing the value

$$B(x) = \frac{\mu_0}{4\pi} \cdot \int_{\Omega} j \times \frac{x-y}{\|x-y\|^3} dV(y) = \frac{\mu_0}{4\pi} \cdot \int_{\Omega} j^P \times \frac{x-y}{\|x-y\|^3} dV(y) - \frac{\mu_0}{4\pi} \cdot \int_{\Omega} \sigma \nabla u \times \frac{x-y}{\|x-y\|^3} dV(y)$$

for all positions $x \notin \Omega$ of interest. We call

$$B^P(x) := \frac{\mu_0}{4\pi} \cdot \int_{\Omega} j^P \times \frac{x-y}{\|x-y\|^3} dV(y)$$

the *primary magnetic field* and

$$B^S(x) := -\frac{\mu_0}{4\pi} \cdot \int_{\Omega} \sigma \nabla u \times \frac{x-y}{\|x-y\|^3} dV(y)$$

the *secondary magnetic field*. Note that we have $B = B^P + B^S$.

We in particular see that if we want to solve the MEG forward problem by evaluating the Biot-Savart integral, we first need to solve the EEG forward problem in order to have access to the electric potential u .

2.2 The Stimulation Forward Problems

In the previous section, we were interested in modeling the process of measuring the electromagnetic fields generated by neural activity. In this section, we are instead interested in modeling the “opposite” approaches of externally applying electromagnetic fields with the goal of modifying neural activity.

One such approach that is of particular interest to this thesis is the so-called *transcranial electric stimulation*, or *TES*. Here, one applies a current to the boundary of the head, which induces an electric field inside the head. It has long been known that this can change the excitability of the affected cortex [139]. Later on, it has also been shown that electric stimulation can have a beneficial effect on a variety of neurological and psychiatric conditions, such as depression [24] or stroke [93]. In recent years, there have furthermore been a number of publications showing that transcranial direct current stimulation can reduce the epileptic activity in drug-resistant epilepsy patients [104, 7, 241], and even a company offering epicranially transplanted stimulation electrodes for epilepsy patients has been founded [187]. In chapter 4 of this thesis, we will present our own stimulation study, where we performed a mathematically optimized electric brain stimulation on an epilepsy patient with the goal of reducing epileptic activity.

From a modeling perspective, one proceeds for the TES in a similar manner as for the EEG problem in section 2.1, see e.g. [224]. First, one again assumes a quasi-static approximation of Maxwell’s equations. Note that the quasi-static approximation assumes $\frac{\partial E}{\partial t} = 0$, which when applied to (2.4) implies

$$\int_{\partial\Omega} \langle j, \eta \rangle dS = \int_{\Omega} \operatorname{div}(j) dV = \int_{\Omega} \frac{1}{\mu_0} \cdot \operatorname{div}(\operatorname{rot}(B)) dV = 0.$$

Thus, for the quasi-static approximation to be valid, the total current injected into the head must equal the amount of current extracted from the head.

In contrast to the EEG case, we now assume that the background neural activity is negligible, i.e., $j^P = 0$.

In the PEM model, we then assume that we have $L \in \mathbb{N}$ electrodes at the positions $x_1, \dots, x_L \in \partial\Omega$. At the l -th electrode position, we inject a current of strength $I_l \in \mathbb{R}$ into the head, where we assume that the sum of I_1, \dots, I_L is zero. At the remaining positions on the boundary, we still model zero current outflow. This leads to the following problem formulation.

Definition 2.2.1 (tDCS forward problem, PEM model). Let Ω denote the head domain with outer normal η and σ the electrical conductivity. Assume that we have $L \in \mathbb{N}$ electrodes at positions $x_1, \dots, x_L \in \partial\Omega$, and that at the l -th electrode we inject a current of strength $I_l \in \mathbb{R}$. Furthermore assume that $\sum_{l=1}^L I_l = 0$. Then, the electric potential $u : \Omega \rightarrow \mathbb{R}$ in the head due to this current injection is given by

$$\begin{aligned} \operatorname{div}(\sigma \nabla u) &= 0 && \text{in } \Omega \\ \langle \sigma \nabla u, \eta \rangle &= \sum_{l=1}^L I_l \cdot \delta_{x_l} && \text{on } \partial\Omega, \end{aligned}$$

where δ_{x_l} denotes the Dirac distribution at x_l .

Note that in the definition above, a value $I_l > 0$ is interpreted as current flowing *into the head*, and a value $I_l < 0$ as a current flowing *out of the head*. An electrode where $I_l > 0$ is called an *anode*, and an electrode where $I_l < 0$ is called a *cathode*. This is the typical convention in the brain stimulation community [131]. Remembering that the current density is given by $-\sigma \nabla u$ and that η points *out of the head*, one can quickly check that this convention indeed leads to definition 2.2.1.

For the CEM model the only difference to the derivation in section 2.1 is that we now model $j^P = 0$, and the current over the l -th electrode interface e_l is now not assumed to be 0, but instead some value $I_l \in \mathbb{R}$. This leads to the following definition.

Definition 2.2.2 (tDCS forward problem, CEM model). Let Ω denote the head domain with outer normal η . Let σ denote the electrical conductivity. Assume that we employ L electrodes with pairwise disjoint contact interfaces $e_1, \dots, e_L \subset \partial\Omega$ and effective contact impedances $z_1, \dots, z_L \in \mathbb{R}$. Furthermore, assume that at the l -th electrode we inject a current of strength $I_l \in \mathbb{R}$, and that we have $\sum_{l=1}^L I_l = 0$. Then, the electric potential $u : \Omega \rightarrow \mathbb{R}$ in the head and the potentials $U_1, \dots, U_L \in \mathbb{R}$ in the electrodes due to this current injection are given by

$$\begin{aligned} \operatorname{div}(\sigma \nabla u) &= 0 && \text{in } \Omega \\ \langle \sigma \nabla u, \eta \rangle &= 0 && \text{on } \partial\Omega \setminus \bigcup_{l=1}^L e_l \\ z_l \cdot \langle \sigma \nabla u, \eta \rangle &= U_l - u && \text{on } e_l \text{ for } 1 \leq l \leq L \\ \int_{e_l} \langle \sigma \nabla u, \eta \rangle dS &= I_l && \text{for } 1 \leq l \leq L \end{aligned}$$

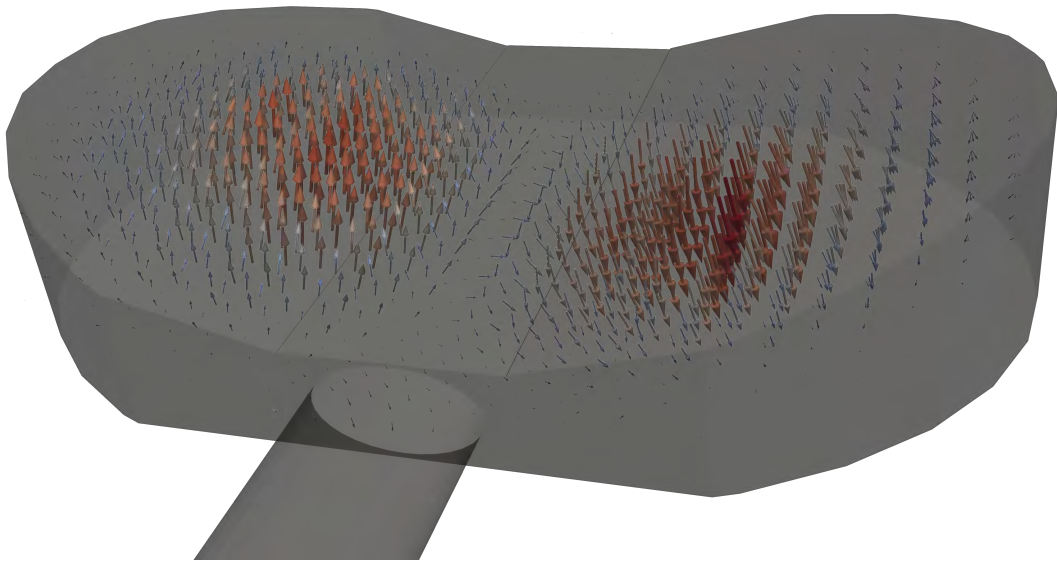
As a final stimulation modality, we want to introduce the so-called *transcranial magnetic stimulation*, or *TMS* for short. In TMS, one places a coil close to the head of the subject and applies a time-varying current through the coil. As described by Maxwell's equations, this time-varying current then produces a time-varying electric field, which can have an effect on the neural activity of the subject. Typically, the coil is built in a figure-8 shape, consisting of two halves with opposite winding. This produces a focal current directly below the center of the coil. A typical TMS setup is illustrated in figure 2.1a, and the corresponding electric field is illustrated in figure 2.2.

The first step towards modelling TMS is again a simplification of Maxwell's equations. The standard reference is [226]. There, the authors first argue that the term $\epsilon_0 \frac{\partial E}{\partial t}$ in (2.4) is negligible. Then, we again have

$$0 = \operatorname{div}(\operatorname{rot}(B)) = \mu_0 \operatorname{div}(j) \iff \operatorname{div}(j) = 0.$$

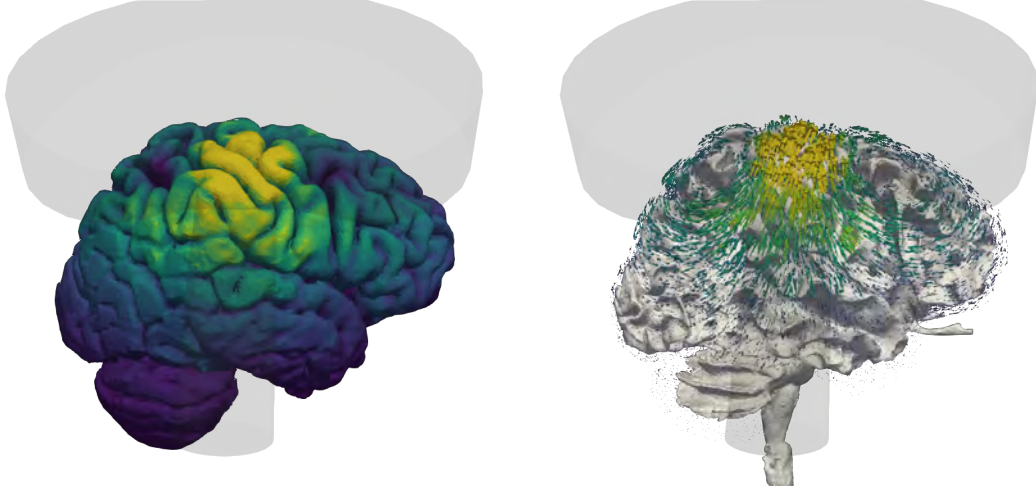


(a) TMS coil placed roughly over the motor cortex. The volume conductor model was created from T1 and T2 MRI images using the CHARM pipeline [158], and white matter conductivity tensors were estimated from DTI data [204]. The MRI scans were performed on a healthy adult male subject who gave written informed consent according to the declaration of Helsinki prior to the measurement, which was approved by the ethics committee of the University of Münster (#2015-263-f-S). The coil is modelled after MagVenture's MC-B70 coil (see e.g. <https://magventure.com/de/products/mc-b70/>), which is one of the two coils currently owned by the IBB, using the geometry description from [47].



(b) Magnetic dipole approximation of the TMS coil from figure 2.1a. The dipole approximation is due to [47] and [123]. The magnetic dipoles are colored and scaled according to the norm of their moments. Note how the opposite winding direction of the two halves of the figure-8 coil is reflected in the opposite orientation of the approximating magnetic dipoles.

Figure 2.1: Typical TMS setup



(a) Magnitude of the electric field plotted on the cortex.

(b) Electric field vectors sampled in the cortex, overlaid on the white matter. The field vectors are scaled and colored according to their magnitudes.

Figure 2.2: Simulated electric field induced by the TMS stimulation setup from figure 2.1, for a rate of change of the current of $\frac{dI}{dt} = -1$. The simulation was performed by interpreting each magnetic dipole from figure 2.1b as an MEG coil and direction pair, computing the corresponding MEG transfer matrix using DUNEuro, summing up the rows of the transfer matrix, and interpreting the resulting vector as the coefficients of a FEM trial function. Finally, one then adds the negative gradient of this FEM trial function to the sum of the magnetic dipole vector potentials. We refer to section 2.9.4 for a more detailed explanation on why this works. This is, admittedly, a rather convoluted and inefficient way of performing this simulation, but it serves as a proof of concept that a toolbox capable of MEG simulations is also capable of TMS simulations, by virtue of reciprocity laws.

Furthermore, following the same argument as given immediately after (2.6), we can again derive that the Biot-Savart law

$$B(x, t) = \frac{\mu_0}{4\pi} \cdot \int_{\mathbb{R}^3} j(y, t) \times \frac{x - y}{\|x - y\|^3} dV(y)$$

is valid, where we now explicitly denote the time dependence, as this will be of fundamental importance later on. We now define

$$A(x, t) := \frac{\mu_0}{4\pi} \cdot \int_{\mathbb{R}^3} \frac{j(y, t)}{\|x - y\|} dV(y) \stackrel{y=z+x}{=} \frac{\mu_0}{4\pi} \cdot \int_{\mathbb{R}^3} \frac{j(z + x, t)}{\|z\|} dV(z). \quad (2.8)$$

Then, we have

$$\operatorname{div}(A)(x, t) = \frac{\mu_0}{4\pi} \int_{\mathbb{R}^3} \frac{\operatorname{div}(j)(z + x, t)}{\|z\|} dV(z) \stackrel{\operatorname{div}(j)=0}{=} 0,$$

and applying (2.5) to the scalar field $x \mapsto \frac{1}{\|x - y\|}$ and the (constant) vector field $x \mapsto j(y, t)$

yields

$$\begin{aligned}\operatorname{rot}(A)(x, t) &= \frac{\mu_0}{4\pi} \cdot \int_{\mathbb{R}^3} \nabla_x \left(\frac{1}{\|x - y\|} \right) \times j(y, t) dV(y) \\ &= \frac{\mu_0}{4\pi} \cdot \int_{\mathbb{R}^3} j(y, t) \times \frac{x - y}{\|x - y\|^3} dV(y) \\ &= B(x, t).\end{aligned}$$

Written more succinctly, if we define A by (2.8), we have $\operatorname{div}(A) = 0$ and $\operatorname{rot}(A) = B$. The vector field A is called the *magnetic vector potential*. Then, equation (2.2) implies that

$$\operatorname{rot}(E) = -\frac{\partial B}{\partial t} \iff \operatorname{rot}\left(E + \frac{\partial A}{\partial t}\right) = 0,$$

and we can hence write

$$E = -\frac{\partial A}{\partial t} - \nabla u \tag{2.9}$$

for a scalar potential u . Here, one calls $E^P := -\frac{\partial A}{\partial t}$ the *primary electric field* and $E^S := -\nabla u$ the *secondary electric field*. Now, we assume that the background activity of the brain is negligible, and the current density in the head is thus given by $j = \sigma E$. This then implies

$$0 = \operatorname{div}(j) \iff \operatorname{div}(\sigma \nabla u) = -\operatorname{div}\left(\sigma \frac{\partial A}{\partial t}\right).$$

We now again assume zero current outflow at the boundary, which implies

$$0 = \langle j, \eta \rangle \iff \langle \sigma \nabla u, \eta \rangle = -\left\langle \sigma \frac{\partial A}{\partial t}, \eta \right\rangle$$

on $\partial\Omega$, where η again denotes the outer normal of the head domain.

Finally, looking at (2.8), one sees that while the current in the TMS coil is driving the stimulation, the current induced in the brain has itself an influence on the magnetic vector potential. The discussion in [226], however, shows that the influence of the currents inside the head on the vector potential A is negligible, and one hence models the vector potential A as arising solely from the TMS stimulation coil. Then, one can model the magnetic vector potential as

$$A(x, t) = f(t) \cdot A^s(x),$$

where A^s is the vector potential generated by the coil at some baseline current strength, e.g. 1 A, and f is the timecourse that scales the current strength from 1 A to the current strength at timepoint t .

Definition 2.2.3 (TMS forward problem). Assume we perform a TMS stimulation with a coil generating a magnetic vector potential $A(x, t) = f(t) \cdot A^s(x)$. Then, the TMS forward problem consists of computing the field

$$E = -f'(t) \cdot A^s(x) - \nabla u$$

on the head domain Ω . If η denotes the outer normal of Ω , the scalar potential u is described by

$$\begin{aligned}\operatorname{div}(\sigma \nabla u) &= -f'(t) \cdot \operatorname{div}(\sigma A^s) && \text{in } \Omega \\ \langle \sigma \nabla u, \eta \rangle &= -f'(t) \cdot \langle \sigma A^s, \eta \rangle && \text{on } \partial\Omega\end{aligned}$$

This leaves open the question of how to compute A , or respectively A^S . One strategy consists of directly evaluating the integral (2.8), by, e.g., creating a model of the geometry of the coil and performing numerical integration over this coil model. This is the approach that much of the older literature takes, see e.g. [169]. More recently, papers have pointed out that it is quite challenging to accurately model the geometry of the TMS stimulation coils, and instead proposed to model the coil by a linear combination of magnetic dipoles, which are fitted based on real measurement data of the fields generated by the coils, see e.g. [47, 123]. This is also the approach taken by SimNIBS, one of the most commonly used toolboxes for TMS modeling.¹ In figure 2.1b we have visualized such an approximation.

Finally, remember that we defined $E^P := -\frac{\partial A}{\partial t}$ and $E^S = -\nabla u$, and that we have $E = E^P + E^S$. Here, the primary field E^P is generally simple to compute, and the secondary field E^S is described by an elliptic partial differential equation and is thus generally more difficult to compute. Note the structural similarity with the primary and secondary fields in the MEG forward problem. In the section on reciprocity laws, we will see that this is not only a superficial similarity. One further similarity is given by the following theorem.

Theorem 2.2.4. *Assume an unbounded, homogeneous, and isotropic conductor, i.e. $\Omega = \mathbb{R}^3$ and $\sigma = c \cdot I_3$, where $0 < c \in \mathbb{R}$ and I_3 is the identity matrix in 3 dimensions. Furthermore, assume that the contribution of $\epsilon_0 \frac{\partial E}{\partial t}$ in (2.4) is negligible. We then have the following.*

- Let j^P denote a compactly supported impressed current density². If $\frac{\partial B}{\partial t}$ in (2.2) is negligible, as we assume in the definition of the MEG forward problem, we have for $x \in \mathbb{R}^3$ that

$$B(x) = B^P(x) := \frac{\mu_0}{4\pi} \cdot \int_{\mathbb{R}^3} j^P(y) \times \frac{x - y}{\|x - y\|^3} dV(y).$$

In particular, we have that the magnetic field contains no contribution due to the passive currents given by σE , i.e., the secondary magnetic field, as defined in 2.1.3, does not contribute to the total magnetic field.

- Assume we perform TMS stimulation with a coil generating a magnetic vector potential A . Then, for $x \in \mathbb{R}^3$ and $t \in \mathbb{R}$ the generated electric field E is given by

$$E(x, t) = E^P(x, t) := -\frac{\partial A}{\partial t}(x, t).$$

In particular, when compared with definition 2.2.3, we see that there is no secondary field contribution.

Proof. Based on physical intuition, we again assume that the electric and magnetic fields decay at infinity, i.e. $B(x) \xrightarrow{x \rightarrow \infty} 0$ and $E(x) \xrightarrow{x \rightarrow \infty} 0$. In the following argument, we will need the fact that a sufficiently regular function $f : \mathbb{R}^3 \rightarrow \mathbb{R}$ decaying at infinity is already uniquely determined by Δf , as is e.g. shown in [64], §16, Corollar 2.

We start with the MEG case. Here, note that since $\frac{\partial B}{\partial t}$ is negligible we have $\text{rot}(E) = 0$ by (2.2). Now, since $\epsilon_0 \frac{\partial E}{\partial t}$ is also negligible and the total current density is given by $j = j^P + \sigma E$, we can

¹see <https://simnibs.github.io/simnibs/build/html/index.html>

²We use the term *impressed current density* instead of *neural activity*, since the term “neural activity” does not really make sense if we assume an unbounded conductor. Note, however, that from a purely electromagnetic view, we still interpret this current in the same way, i.e., j^P is some externally driven current, which generates an electric field that in turn generates passive return currents. We thus still model the total current density as $j = j^P + \sigma E$.

compute

$$-\Delta(B) \stackrel{(2.3)}{=} \nabla(\operatorname{div}(B)) - \Delta(B) \stackrel{(2.6)}{=} \operatorname{rot}(\operatorname{rot}(B)) \stackrel{\text{quasi-static (2.4)}}{=} \mu_0 \operatorname{rot}(j^P) + \mu_0 \operatorname{rot}(\sigma E) \\ \stackrel{\sigma \text{ constant and isotropic}}{=} \mu_0 \operatorname{rot}(j^P) + \mu_0 \cdot c \cdot \operatorname{rot}(E) \stackrel{\operatorname{rot}(E)=0}{=} \mu_0 \operatorname{rot}(j^P).$$

Continuing with the argument in the same way as after deriving (2.7), we arrive at $B(x) = B^P(x)$.

For the TMS case, we have no impressed current, i.e., we model $j = \sigma E$. But then we have

$$0 = \operatorname{div}(\operatorname{rot}(B)) = \mu_0 \operatorname{div}(j) \stackrel{\sigma \text{ constant and isotropic}}{=} \mu_0 \cdot c \cdot \operatorname{div}(E) \iff \operatorname{div}(E) = 0.$$

Now, since both E and A are divergence-free, we can compute similarly as before that

$$-\Delta(E) = \operatorname{rot}(\operatorname{rot}(E)) \stackrel{(2.2)}{=} \operatorname{rot}(\operatorname{rot}(-\frac{\partial A}{\partial t})) = \Delta(\frac{\partial A}{\partial t}).$$

Looking at (2.8), one can see that A decays at infinity. Hence, we have $E = -\frac{\partial A}{\partial t} = E^P$. \square

We thus see that in both MEG and TMS, we have a decomposition into a primary and a secondary field, where the primary field is the field that would arise in an unbounded isotropic homogeneous volume conductor. In particular, one could say that the secondary fields arise due to tissue inhomogeneities.

For the TMS case in particular, the combination of $\operatorname{div}(A) = 0$ and (2.1) implies for the charge density ρ that

$$\frac{\rho}{\epsilon_0} = \operatorname{div}(E) = \operatorname{div}(-\frac{\partial A}{\partial t} - \nabla u) = -\Delta u.$$

If now $U \subset \Omega$ is an open set such that $\sigma|_U$ is constant and isotropic, the quasi-static approximation of (2.4) implies that on U we have

$$0 = \operatorname{div}(j) = \operatorname{div}(\sigma E) = \sigma \operatorname{div}(-\frac{\partial A}{\partial t} - \nabla u) = -\sigma \Delta u = -\frac{\sigma}{\epsilon_0} \rho \iff \rho = 0.$$

These two observations together, and in comparison with theorem 2.2.4, motivate the statement that the secondary electric field encountered in the TMS stimulation of an inhomogeneous tissue arises due to the accumulation of surface charges at tissue interfaces. This statement is commonly encountered, but not explained, in the TMS literature, see e.g. [202].

In the EEG and MEG case, equation (2.1) again implies $-\Delta u = \frac{\rho}{\epsilon_0}$, and by almost the same computation as above, we have $\rho = 0$ in regions where σ is constant and isotropic and where $j^P = 0$. One could thus, similarly to the TMS case, claim that the electric potential, and then also the secondary magnetic field, are due to charge accumulation at the tissue interfaces and regions of non-zero neural activity. This perspective is, however, to the best of our knowledge, not really present in the EEG and MEG literature. One notable recent exception is [124], where the authors use this perspective to derive a boundary element method for the EEG and MEG forward problems, which was originally thought to be a new approach. It, however, later turned out that, while the derivation in [124] is, to the best of our knowledge, new, the numerical approach it yields is the same as the so-called *single layer BEM approach*, which had already been introduced to the MEEG community in [110].

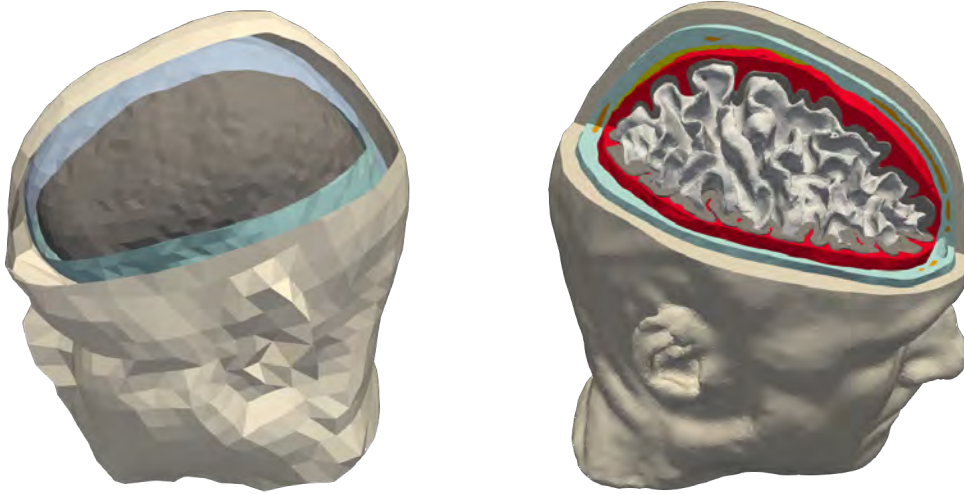
2.3 Choosing an approach for solving the forward problems

In bioelectromagnetism, one is interested in solving the forward problems defined in the preceding sections. In simple geometries, such as spheres with radially symmetric conductivity profiles, in combination with simple source terms for $\frac{\partial A}{\partial t}$ in the TMS case and j^P in the EEG/MEG case, one can derive analytical expressions for the solutions of the forward problems. In particular for spherical geometries, one can look at [236, 39] for analytical PEM EEG formulas, [61, 62] for analytical PEM tDCS formulas, [178] for analytical MEG formulas, and [98] for analytical TMS formulas. Even though there are, to this day, many publications that use a spherical model for their investigations, it is evident that the head is not a sphere, and it has been shown that ignoring this can introduce serious errors. One such example is [202], where the authors show that a spherical head model strongly underestimates the field strengths of TMS stimulations in gyral crowns, and further illustrate how the complex folding of the cortex introduces considerable deviations of the electric field in comparison to a spherical model. Note that by virtue of the reciprocity laws, which we will discuss in section 2.9 of this thesis, this result directly carries over to a pendant for MEG modeling. The importance of a realistic geometry modeling has also been demonstrated for the EEG problem, see e.g. [212, 216], which only becomes more pronounced in the presence of non-standard head geometries, such as skulls containing holes [147, 117], which can e.g. be encountered in epilepsy patients that underwent some form of invasive surgery.

For head models that are more realistic than an ellipsoid, there are, to the best of our knowledge, no known analytical solutions for the forward problems. If one decides to pursue a more realistic modeling approach, one thus has to employ some numerical method for the solution of the respective forward problem. Over the years, numerous such approaches have been suggested, see e.g. [140] or [36]. The two most commonly used numerical approaches, however, are the *boundary element method (BEM)* and the *finite element method (FEM)*. We want to note that while both BEM and FEM approaches are used for all of the previously derived forward problems, the MEEG community generally favors BEM approaches, while the brain stimulation community generally favors FEM approaches. There is, to the best of our knowledge, no deeper reason for this, in particular as the stimulation and measurement problems are essentially the same, as we will show in our discussion on reciprocity laws in section 2.9.

When employing a BEM approach, the current default in the bioelectromagnetism community is to model the head as consisting of three layers, namely scalp, skull, and “brain”. Here, “brain” means a homogenized interior compartment encompassing the different brain tissues as well as the CSF. Furthermore, the “skull” compartment merges the compact bone tissue and the spongy bone tissue. Although rarely done, it is of course also possible to model more than 3 tissue compartments in BEM, see e.g. [144]. For FEM approaches, there is not really a “standard” way of building models. Our own previous in-house pipeline [54] defaulted to segmenting 6 tissues, namely white matter, gray matter, CSF, compact bone, spongy bone, and scalp. The CHARM pipeline [158], on the other hand, additionally segments the eyeballs, blood, and muscle tissue. One could, in principle, add even more compartments, as the “true” number of tissues in the human head is larger than 9, but due to uncertainties in the segmentation and the assignment of conductivity values, this is seldom done. Examples of standard BEM and FEM models are visualized in figure 2.3.

If one decides to employ a three layer head model, it has been shown that BEM and FEM lead to a comparable accuracy, with their difference being lower than the error introduced by not modeling the CSF [217], using the at the time most advanced BEM [110] and FEM [116]



(a) Three-compartment BEM model, with surfaces representing the scalp, skull, and “brain”, i.e., everything inside the skull. The model was created from a T1 MRI image using FieldTrip [146], following their tutorial on BEM head model creation at https://www.fieldtriptoolbox.org/tutorial/source/headmodel_eeg_bem/. The underlying MRI image was the same as the one used for figure 2.1a. The brain surface consists of 3k nodes, the skull surface of 2k nodes, and the scalp surface of 1k nodes. Each surface consists of ca. twice as many triangles as it has nodes.

(b) FEM model from figure 2.1a, with parts of the model clipped off to illustrate the embedding of the different segmented tissues inside the head. The model consists of ca. 720k nodes and 4.04 million tetrahedral elements. The white matter is colored white, the grey matter is colored grey, and the CSF is colored red. Note in particular the complex folding of the grey matter surface, and the corresponding CSF-filled sulci, which are not included in the BEM model in figure 2.3a.

Figure 2.3: Comparison of BEM and FEM head models of the same subject. Both models were created using the default parameters of the underlying toolboxes.

approaches. Since then, there have been significant improvements on the FEM side [218, 18, 91], and it might hence be interesting to revisit this comparison at some point. However, our primary conclusion from [217] is the importance of modeling the CSF. This is furthermore complemented by [216] and [237], where the authors demonstrate that differentiating white matter and gray matter also has a strong influence on the forward simulation, and furthermore that modelling the white matter conductivity as an anisotropic tensor instead of an isotropic scalar value also has a noticeable impact. Given these studies, we argue that if one aims to build a realistic head model, one should include the distinction between CSF, gray matter, and white matter, and, if available, one should also include DTI-derived anisotropic white matter conductivity tensors. This, in turn, incentivises the use of FEM approaches, as including complicated tissue segmentations as well as conductivity anisotropy poses no real additional challenge for FEM, which, in contrast, cannot be said about BEM approaches.

As already mentioned above, we want to emphasize that there is no principal reason that prevents BEM methods from including multiple interfaces with a complex folding. Nevertheless, if one tries

to push the method to use more realistic segmentations, one encounters a number of difficulties. First of all, BEM approaches lead to linear equations $A \cdot x = b$ with a dense matrix A , whose number of rows depends on the surface mesh resolution. Using naive solver strategies, this quickly becomes computationally infeasible. This can, however, be overcome by using the *fast multipole method (FMM)*, which enables an evaluation of products of the form $A \cdot x$ without needing to explicitly assemble the matrix A . Using this strategy, BEM approaches can also be applied to high-resolution meshes [111], though this comes at the cost of a significantly increased complexity of the implementation. Additionally, while BEM approaches can be extended to cover anisotropic conductivity tensors [156, 145], this is a highly nontrivial generalization from the isotropic case, and thus adds another layer of difficulty to a potential implementation.

As we explicitly want to employ realistic and anisotropic head models, we will thus focus on finite element approaches in this thesis. Accordingly, we will now discuss how to derive FEM formulations for the forward problems that were introduced in sections 2.1 and 2.2.

2.4 Deriving FEM formulations for the stimulation forward problems

Our new contributions are mainly concerned with the EEG and MEG forward problems. In this section, we will thus first present a quick derivation of the weak formulations of the tES and TMS forward problems, with the goal of preparing the later discussion on reciprocity, and then primarily focus on a more detailed discussion of the EEG and MEG problems starting in section 2.5.

In the following, we will always write $\Omega \subset \mathbb{R}^3$ for the compact head domain with outer normal η , and σ for the potentially anisotropic conductivity profile of the head. Furthermore, we assume that there are constants $0 < c, C \in \mathbb{R}$ such that $0 < c \cdot I \leq \sigma \leq C \cdot I$ on Ω , where I denotes the identity matrix.³

2.4.1 The tDCS forward problem

We start with the tDCS forward problem, see definitions 2.2.1 and 2.2.2. Let $I_1, \dots, I_L \in \mathbb{R}$ denote a current injection pattern. For a suitable trial function u and test function φ , we can compute in both the PEM and CEM cases that

$$\begin{aligned} 0 &= \int_{\Omega} \operatorname{div}(\sigma \nabla u) \varphi \, dV = \int_{\partial\Omega} \langle \sigma \nabla u, \eta \rangle \varphi \, dS - \int_{\Omega} \langle \sigma \nabla u, \nabla \varphi \rangle \, dV \\ &\iff \int_{\Omega} \langle \sigma \nabla u, \nabla \varphi \rangle \, dV = \int_{\partial\Omega} \langle \sigma \nabla u, \eta \rangle \varphi \, dS. \end{aligned} \quad (2.10)$$

In the PEM case, we get

$$\int_{\partial\Omega} \langle \sigma \nabla u, \eta \rangle \varphi \, dS = \sum_{i=1}^L I_i \cdot \varphi(x_i),$$

³In our concrete application c can e.g. be chosen as the skull conductivity and C as the CSF conductivity.

while in the CEM case, we get for an arbitrary vector $c = (c_1, \dots, c_L) \in \mathbb{R}^L$ that

$$\begin{aligned} \int_{\partial\Omega} \langle \sigma \nabla u, \eta \rangle \varphi \, dS &= \sum_{l=1}^L \int_{e_l} \langle \sigma \nabla u, \eta \rangle \varphi \, dS = \sum_{l=1}^L \int_{e_l} \langle \sigma \nabla u, \eta \rangle (\varphi - c_l) \, dS + c_l \cdot I_l \\ &= \sum_{l=1}^L \frac{1}{z_l} \cdot \int_{e_l} (U_l - u)(\varphi - c_l) \, dS + c_l \cdot I_l, \end{aligned}$$

which in combination with (2.10) leads to

$$\int_{\Omega} \langle \sigma \nabla u, \nabla \varphi \rangle \, dV + \sum_{l=1}^L \frac{1}{z_l} \cdot \int_{e_l} (u - U_l)(\varphi - c_l) \, dS = \sum_{l=1}^L c_l \cdot I_l.$$

This essentially finishes the derivation. The “only” thing missing is specifying spaces for the trial and test functions u and φ . While this can be done in a standard way for the CEM formulation, the PEM model is more technically involved due to the Dirac distributions on the boundary. We refer to [85, 120] for details.

Definition 2.4.1 (Weak tDCS forward problem, PEM model). Let $I_1, \dots, I_L \in \mathbb{R}$ with $\sum_{l=1}^L I_l = 0$ denote a current injection pattern. Then, after suitably defining a trial space U and a test space V , the weak formulation of the tDCS PEM forward problem is given by

$$\text{Find } u \in U \text{ such that } \int_{\Omega} \langle \sigma \nabla u, \nabla v \rangle \, dV = \sum_{i=1}^L I_i \cdot v(x_i) \text{ for all } v \in V. \quad (2.11)$$

Note that in our FEM discretizations, we always employ function spaces that consist of continuous functions and where pointwise evaluation is hence well-defined. As a consequence, we have, in this thesis, always chosen the same trial and test spaces for numerical PEM tDCS simulations.

In the following, let $H^1(\Omega)$ denote the Sobolev space of square integrable functions with square integrable weak derivatives.

Definition 2.4.2 (Weak tDCS forward problem, CEM model). Assume we have L electrodes with contact interfaces $e_1, \dots, e_L \subset \partial\Omega$ and contact impedances $0 < z_1, \dots, z_L \in \mathbb{R}$. Let $I_1, \dots, I_L \in \mathbb{R}$ with $\sum_{l=1}^L I_l = 0$ denote a current injection pattern. Furthermore, let $\mathcal{H} := H^1(\Omega) \oplus \mathbb{R}^L$. Then, the weak formulation of the tDCS CEM forward problem is given by

Find $(u, U) \in \mathcal{H}$ such that

$$\int_{\Omega} \langle \sigma \nabla u, \nabla v \rangle \, dV + \sum_{l=1}^L \frac{1}{z_l} \cdot \int_{e_l} (u - U_l)(v - V_l) \, dS = \sum_{l=1}^L V_l \cdot I_l \quad \text{for all } (v, V) \in \mathcal{H}.$$

Because the bilinear forms used in definitions 2.4.1 and 2.4.2 arise multiple times in the following, we want to quickly introduce some notation and derive some properties. In the following, we denote for $u, v \in H^1(\Omega)$

$$a_{\text{PEM}}(u, v) := \int_{\Omega} \langle \sigma \nabla u, \nabla v \rangle \, dV \quad (2.12)$$

and for $U, V \in \mathbb{R}^L$

$$a_{\text{CEM}}((u, U), (v, V)) := a_{\text{PEM}}(u, v) + \sum_{l=1}^L \frac{1}{z_l} \cdot \int_{e_l} (u - U_l)(v - V_l) dS. \quad (2.13)$$

We then have the following results, whose proofs can e.g. be found in [239] and [196]. Because we want to present a relatively self-contained derivation of our new contributions, we will present sketches of these proofs.

Lemma 2.4.3. *a_{PEM} is bilinear, symmetric, and bounded on $H^1(\Omega)$, and a_{CEM} is bilinear, symmetric and bounded on \mathcal{H} . Furthermore, let*

$$H_*^1(\Omega) := \{u \in H^1(\Omega) \mid \int_{\Omega} u dV = 0\}$$

and

$$\mathcal{H}_* := \{(u, U) \in \mathcal{H} \mid \int_{\Omega} u dV = -\sum_{l=1}^L U_l\}.$$

Then $H_*^1(\Omega)$ is a closed subspace of $H^1(\Omega)$, \mathcal{H}_* is a closed subspace of \mathcal{H} , a_{PEM} is coercive on $H_*^1(\Omega)$, and a_{CEM} is coercive on \mathcal{H}_* .

Proof. First, note that $H_*^1(\Omega)$ is the orthogonal complement of the constant 1 function, and \mathcal{H}_* is the orthogonal complement of $(1, 1)$. Hence, both of these spaces are closed subspaces. Furthermore, Poincaré's inequality (see e.g. [57], Chapter 5.8) implies that for $u \in H_*^1(\Omega)$ we have

$$\begin{aligned} \|u\|_{H^1(\Omega)}^2 &= \|u - \frac{1}{\text{vol}(\Omega)} \cdot \int_{\Omega} u dV\|_{H^1(\Omega)}^2 \leq C_1 \cdot \|\nabla u\|_{L^2(\Omega)}^2 \leq C_2 \cdot \int_{\Omega} \langle \sigma \nabla u, \nabla u \rangle dV \\ &= C_2 \cdot a_{\text{PEM}}(u, u), \end{aligned}$$

where C_2 depends only on Ω and σ . Thus, a_{PEM} is coercive on $H_*^1(\Omega)$.

Now, let $(u, U) \in \mathcal{H}_*$. Then, (u, U) is orthogonal to (c, c) for an arbitrary $c \in \mathbb{R}$, and hence by the Pythagorean theorem

$$\|(u, U)\|_{\mathcal{H}}^2 \leq \|(u + c, U + c)\|_{\mathcal{H}}^2 = \|u + c\|_{H^1(\Omega)}^2 + \|U + c\|_2^2.$$

A straightforward computation reveals that for $a, b \in \mathbb{R}$ we have $(a + b)^2 \leq 2a^2 + 2b^2$. For the second term, we can now compute

$$\begin{aligned} \|U + c\|_2^2 &= \sum_{l=1}^L (U_l + c)^2 = \sum_{l=1}^L \frac{1}{\text{area}(e_l)} \cdot \int_{e_l} (U_l + c)^2 dS \\ &\leq \sum_{l=1}^L \frac{2}{\text{area}(e_l)} \cdot \int_{e_l} (u + c)^2 dS + \frac{2}{\text{area}(e_l)} \cdot \int_{e_l} (U_l - u)^2 dS \\ &\leq \frac{2}{\min_l \text{area}(e_l)} \cdot \int_{\bigcup_l e_l} (u + c)^2 dS + \frac{2 \max_l z_l}{\min_l \text{area}(e_l)} \cdot \sum_{l=1}^L \frac{1}{z_l} \cdot \int_{e_l} (u - U_l)^2 dS. \end{aligned}$$

Now, by the trace theorem (see e.g. [57], Chapter 5.5) we have

$$\int_{\bigcup_l e_l} (u + c)^2 dS \leq \|u + c\|_{L^2(\Omega)}^2 \leq C \cdot \|u + c\|_{H^1(\Omega)}^2,$$

which, in combination with the preceding computation, yields

$$\|(u, U)\|_{\mathcal{H}}^2 \leq C \cdot \left(\|u + c\|_{H^1(\Omega)}^2 + \sum_{l=1}^L \frac{1}{z_l} \cdot \int_{e_l} (u - U_l)^2 dS \right).$$

Applying this to $c_0 := -\frac{\int_{\Omega} u dV}{\text{vol}(\Omega)}$, we can again apply Poincaré's inequality to obtain

$$\|u + c_0\|_{H^1(\Omega)}^2 \leq C \int_{\Omega} \langle \sigma \nabla u, \nabla u \rangle dV,$$

with the same computation as in the PEM case. Thus we have

$$\epsilon \cdot \|(u, U)\|_{\mathcal{H}}^2 \leq a_{\text{CEM}}((u, U), (u, U))$$

with $0 < \epsilon$ independent of (u, U) , i.e. a_{CEM} is coercive on \mathcal{H}_* .

The bilinearity and symmetry of a_{PEM} and a_{CEM} follow immediately from the definition. The boundedness of a_{PEM} directly follows from the Cauchy-Schwarz inequality. It thus only remains to show the boundedness of a_{CEM} . Since we already know that a_{PEM} is bounded on $H^1(\Omega)$, it suffices to investigate the second term in (2.13). Then, the Cauchy-Schwarz inequality implies for $(u, U), (v, V) \in \mathcal{H}$ that

$$\begin{aligned} \sum_{l=1}^L \frac{1}{z_l} \int_{e_l} (u - U_l)(v - V_l) dS &\stackrel{\text{Cauchy-Schwarz}}{\leq} C \cdot \sum_{l=1}^L \sqrt{\int_{e_l} (u - U_l)^2 dS} \cdot \sqrt{\int_{e_l} (v - V_l)^2 dS} \\ &\stackrel{\text{Cauchy-Schwarz}}{\leq} C \cdot \sqrt{\sum_{l=1}^L \int_{e_l} (u - U_l)^2 dS} \cdot \sqrt{\sum_{l=1}^L \int_{e_l} (v - V_l)^2 dS}. \end{aligned}$$

Now, we have

$$\begin{aligned} \sum_{l=1}^L \int_{e_l} (u - U_l)^2 dS &\leq \sum_{l=1}^L \int_{e_l} 2 \cdot u^2 + 2 \cdot U_l^2 dS = 2 \cdot \int_{\bigcup e_l} u^2 dS + \sum_{l=1}^L 2 \cdot \text{area}(e_l) \cdot U_l^2 \\ &\leq C \left(\|u\|_{H^1(\Omega)}^2 + \|U\|_2^2 \right) = C \|(u, U)\|_{\mathcal{H}}^2, \end{aligned}$$

where we again used the trace theorem. The same computation also applies to (v, V) , which together with the preceding computation implies the boundedness of a_{CEM} on \mathcal{H} . \square

Remark 2.4.4. On first sight, the argument above might look somewhat different from the one presented in [196], where the space \mathcal{H}_* is not introduced, and all CEM results are framed for the quotient space $\mathcal{H}/_{\mathbb{R} \cdot (1,1)}$.⁴ But since for an arbitrary Hilbert space H with closed subspace V , the spaces V^\perp and $\mathcal{H}/_V$ are isometrically isomorphic via the quotient map, this is really the same as the argument presented above.

⁴Strictly speaking, [196] frame their results over \mathbb{C} . For our purposes, it suffices to work over \mathbb{R} .

Corollary 2.4.5. *For an arbitrary current injection pattern I_1, \dots, I_L with $\sum_{l=1}^L I_l = 0$ the CEM forward problem 2.4.2 has a solution $(u_0, U_0) \in \mathcal{H}$. This solution is unique up to an additive constant, i.e., the set of all solutions is given by $(u_0, U_0) + \mathbb{R} \cdot (1, 1)$.*

Proof. By lemma 2.4.3 it follows that a_{CEM} is bounded and coercive on \mathcal{H}_* . If we now define

$$l : \mathcal{H} \rightarrow \mathbb{R}; (v, V) \mapsto \sum_{l=1}^L I_l \cdot V_l,$$

then $l(1, 1) = \sum_{l=1}^L I_l = 0$, and $|l(v, V)| \leq \|I\|_2 \cdot \|V\|_2 \leq \|I\|_2 \cdot \|(v, V)\|_{\mathcal{H}}$, i.e. l is continuous on \mathcal{H} . But then l is in particular continuous on \mathcal{H}_* , and by the Lax-Milgram theorem (see e.g. [57], Chapter 6.2) there exists a unique $(u_0, U_0) \in \mathcal{H}_*$ such that

$$a_{\text{CEM}}((u_0, U_0), (v, V)) = l(v, V) \quad \text{for all } (v, V) \in \mathcal{H}_*. \quad (2.14)$$

But if now $(v, V) \in \mathcal{H}$ is arbitrary, it can be decomposed as $(v, V) = (v_0, V_0) + c \cdot (1, 1)$, with $(v_0, V_0) \in \mathcal{H}_*$ and $c \in \mathbb{R}$. Since $a_{\text{CEM}}((u_0, U_0), (1, 1)) = 0$ and $l(1, 1) = 0$, this implies

$$a_{\text{CEM}}((u_0, U_0), (v, V)) = a_{\text{CEM}}((u_0, U_0), (v_0, V_0)) = l(v_0, V_0) = l(v, V),$$

i.e., (u_0, U_0) is a solution of the weak CEM tDCS forward problem. If $(u, U) \in \mathcal{H}$ is another solution, one can find $c \in \mathbb{R}$ such that $(u - c, U - c) \in \mathcal{H}_*$. But since $a_{\text{CEM}}((u - c, U - c), (v, V)) = a_{\text{CEM}}((u, U), (v, V))$ for all $(v, V) \in \mathcal{H}$, it follows that $(u - c, U - c)$ is a solution of (2.14) in \mathcal{H}_* . But since this solution is unique, it follows that $(u, U) = (u_0, U_0) + (c, c)$. Finally, it is straightforward to check that each function of this form is indeed a solution of the problem. \square

Corollary 2.4.6. *Let $l : \mathcal{H} \rightarrow \mathbb{R}$ be a continuous linear functional with $l(1, 1) = 0$. Then, the problem*

$$\text{Find } (u, U) \in \mathcal{H} \text{ such that } a_{\text{CEM}}((u, U), (v, V)) = l(v, V) \text{ for all } (v, V) \in \mathcal{H} \quad (2.15)$$

has a solution $(u_0, U_0) \in \mathcal{H}$, and this solution is unique up to an additive constant, i.e., the set of all solutions is given by $(u_0, U_0) + \mathbb{R} \cdot (1, 1)$.

Proof. This directly follows from the proof of corollary 2.4.5. \square

Given the preceding discussion, a natural question that arises is how the complete electrode model relates to the point electrode model. One answer to this question is given by the result that, as the electrode interfaces e_l get smaller, the complete electrode model converges towards the point electrode model, which was first proven by [84]. We will quickly state the result. First, we need to give a rigorous meaning to the idea of electrodes “getting smaller”. For this, we follow the definition given in [84].

Definition 2.4.7. We assume that we have L electrodes $e_1^h, \dots, e_L^h \subset \partial\Omega$ which are parametrized by $h \in (0, h_0)$. Furthermore, we assume that we have L positions x_1, \dots, x_L with $x_l \in e_l^h$ for all $1 \leq l \leq L$ and $0 < h < h_0$. We now say that the electrodes *shrink towards* the positions x_1, \dots, x_L if there is some convex bounded reference domain $Q \subset \mathbb{R}^2$ and maps X_l^h such that $e_l^h = X_l^h(hQ)$, X_l^h is a diffeomorphism, the gramian of X_l^h is bounded from above and below independent of l and h , $\|X_l^h\|_{C^2(\overline{hQ})} \leq C$ independent of l and h , $0 \in hQ$, and $x_l = X_l^h(0)$. We say that the

electrodes *shrink nicely* towards the positions x_1, \dots, x_L if they shrink towards x_1, \dots, x_L and one furthermore has

$$\int_{hQ} x g_l^h(x) dS = 0,$$

where g_l^h denotes the gramian of X_l^h . This last condition means that the parametrizations of the electrodes shrink towards their centers of mass.

Theorem 2.4.8 ([84], Corollary 3.3). *Let $x_1, \dots, x_L \in \partial\Omega$ denote point electrode positions and let $e_1^h, \dots, e_L^h \subset \partial\Omega$ denote electrode interfaces parametrized by $h \in (0, h_0)$. Furthermore, let $I_1, \dots, I_L \in \mathbb{R}$ with $\sum_{l=1}^L I_l = 0$ denote a current injection pattern, and let u_{CEM}^h denote the corresponding CEM tDCS solution and u_{PEM} the corresponding PEM tDCS solution. Finally, let $\Omega_0 \subset \Omega$ be such that $\bar{\Omega}_0 \subset \Omega$. We then have the following.*

- *If the electrodes shrink towards x_1, \dots, x_L , there is a constant $C > 0$, independent of h and I , such that for all $0 < h < h_0$ we have*

$$\|u_{CEM}^h - u_{PEM}\|_{H^1(\Omega_0)/_{\mathbb{R},1}} \leq C \cdot h \cdot \|I\|_2.$$

- *If the electrodes shrink nicely towards x_1, \dots, x_L , there is a constant $C > 0$, independent of h and I , such that for all $0 < h < h_0$ we have*

$$\|u_{CEM}^h - u_{PEM}\|_{H^1(\Omega_0)/_{\mathbb{R},1}} \leq C \cdot h^2 \cdot \|I\|_2.$$

In section 2.9.2, we will see that, using reciprocity and elliptic regularity, this result can be extended to the EEG case. Furthermore, we want to point the interested reader to the excellent master's thesis [66], where the proof of theorem 2.4.8, and applications of the result to tDCS montage optimization, are presented and discussed in-depth.⁵

As a final point in this subsection, we want to discuss how to set up FEM simulations for the tDCS forward problems. This is mainly standard, except for the fact that the forward problems, and then also their Galerkin approximations, do not possess unique solutions, but only solutions that are unique up to an additive constant. This will manifest itself in a rank deficiency of the FEM stiffness matrix. We will now discuss how to address this.

Lemma 2.4.9. *Let V denote a finite-dimensional inner product space and $a : V \times V \rightarrow \mathbb{R}$ a symmetric bilinear form. Furthermore, let $v \in V$ be such that for all $u \in V$ we have $a(v, u) = a(u, v) = 0$. Additionally, assume that a is coercive on $\{v\}^\perp$, and that a basis $\mathfrak{B} = \{b_1, \dots, b_N\}$ of V is given such that $v = \sum_{i=1}^N \lambda_i \cdot b_i$ with $\lambda_1 \neq 0$. Then, define $A = (a_{ij})_{1 \leq i, j \leq N}$ by $a_{ij} := a(b_j, b_i)$, and*

$$\hat{A} := \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ 0 & a_{n2} & \dots & a_{nn} \end{pmatrix}.$$

We then have the following.

- \hat{A} is symmetric positive definite.

⁵<https://www.sci.utah.edu/~wolters/PaperWolters/2022/MasterFrank.pdf>

- If $l : V \rightarrow \mathbb{R}$ is a linear functional with $l(v) = 0$, the problem

$$\text{Find } u \in V \text{ such that } a(u, w) = l(w) \text{ for all } w \in V \quad (2.16)$$

has a solution $u_0 \in V$. The set of all solutions of this problem is given by $u_0 + \mathbb{R} \cdot v$. If we define $y = (y_i)_{1 \leq i \leq N} \in \mathbb{R}^N$ by $y_i = l(b_i)$, and set

$$\hat{y} := \begin{pmatrix} 0 \\ y_2 \\ \vdots \\ y_N \end{pmatrix},$$

the coefficients $x \in \mathbb{R}^N$ in the basis \mathfrak{B} of the unique solution of (2.16) whose coefficient of b_1 is 0 is given as the solution of the system $\hat{A} \cdot x = \hat{y}$.

Proof. Let l be a functional with $l(v) = 0$. Because V is finite-dimensional, a is bounded, and all functionals are continuous. Since by assumption a is coercive on $\{v\}^\perp$, the Lax-Milgram theorem implies that there is a unique $u_0 \in \{v\}^\perp$ such that $a(u_0, w) = l(w)$ for all $w \in \{v\}^\perp$. An arbitrary $w \in V$ can then be decomposed as $w = w_0 + \lambda \cdot v$, with $w_0 \in \{v\}^\perp$ and $\lambda \in \mathbb{R}$. But then

$$a(u_0, w) = a(u_0, w_0) = l(w_0) = l(w),$$

i.e. u_0 solves (2.16). If u is another solution, there is $\lambda \in \mathbb{R}$ with $u - \lambda v \in \{v\}^\perp$. We then have for all $w \in \{v\}^\perp$ that $a(u - \lambda v, w) = l(w)$, which implies $u = u_0 + \lambda \cdot v$. Furthermore, it is straightforward to check that each vector of this form is indeed a solution of (2.16).

Given the above representation of v in the basis \mathfrak{B} , it follows that there is a unique solution of (2.16) whose coefficient representation $x \in \mathbb{R}^N$ in the basis \mathfrak{B} fulfills $x_1 = 0$. Now, if we investigate a solution z of the equation $\hat{A} \cdot z = \hat{y}$, we see that the first row implies $z_1 = 0$. For $2 \leq i \leq N$, this observation together with the condition expressed by the i -th row of $\hat{A} \cdot z = \hat{y}$ implies that

$$l(b_i) = \hat{y}_i = \sum_{j=2}^N a(b_j, b_i) \cdot z_j \stackrel{z_1=0}{=} \sum_{j=1}^N a(b_j, b_i) \cdot z_j = a\left(\sum_{j=1}^N z_j \cdot b_j, b_i\right).$$

Additionally, we have

$$\begin{aligned} l(b_1) &= l\left(\frac{1}{\lambda_1} \cdot \left(v - \sum_{i=2}^N \lambda_i \cdot b_i\right)\right) = -\frac{1}{\lambda_1} \cdot \sum_{i=2}^N \lambda_i \cdot l(b_i) = -\frac{1}{\lambda_1} \cdot \sum_{i=2}^N \lambda_i \cdot a\left(\sum_{j=1}^N z_j \cdot b_j, b_i\right) \\ &= a\left(\sum_{j=1}^N z_j \cdot b_j, \frac{1}{\lambda_1} \cdot \left(v - \sum_{i=2}^N \lambda_i \cdot b_i\right)\right) = a\left(\sum_{j=1}^N z_j \cdot b_j, b_1\right). \end{aligned}$$

Since \mathfrak{B} is a basis of V and a is bilinear, this implies that $u := \sum_{j=1}^N z_j b_j$ fulfills $z_1 = 0$ and is a solution of $a(u, w) = l(w)$ for all $w \in V$. But this implies $z = x$, i.e., we have shown that $\hat{A} \cdot z = \hat{y}$ has at most one solution, namely x . On the other hand, we indeed have $\hat{A} \cdot x = \hat{y}$, as a straightforward computation reveals. Note that this, in particular, implies that the matrix \hat{A} is invertible. Furthermore, it is straightforward to see that \hat{A} is symmetric.

To finish the proof, it thus suffices to show that \hat{A} is symmetric positive semidefinite. If we now have $(z_1, \dots, z_N) \in \mathbb{R}^N$, we can compute

$$\langle z, \hat{A}z \rangle = z_1^2 + \sum_{i=2}^N z_i \cdot \sum_{j=2}^N z_j \cdot a(b_j, b_i) = z_1^2 + a\left(\sum_{j=2}^N z_j \cdot b_j, \sum_{i=2}^N z_i \cdot b_i\right).$$

Now, an arbitrary $w \in V$ can be written as $w = w_0 + \lambda v$, with $w_0 \in \{v\}^\perp$ and $\lambda \in \mathbb{R}$. We then have

$$a(w, w) = a(w_0, w_0) \geq 0,$$

since a is by assumption coercive on $\{v\}^\perp$. This then implies that $\langle z, \hat{A}z \rangle \geq 0$, which finishes the proof. \square

Lemma 2.4.9 allows us to express the FEM approximations of the forward problems of bioelectromagnetism as linear equations with a symmetric positive definite matrix. For PEM approaches, one chooses $v = 1$, and for CEM approaches, one chooses $v = (1, 1)$. The FEM trial spaces V_h for the functions on the head domain that are used in this thesis all fulfill $V_h \subset H^1(\Omega) \cap C(\bar{\Omega})$. In the PEM case, we search for solutions directly in V_h , and in the CEM case, we search for solutions in $V_h \oplus \mathbb{R}^L$. The coercivity condition on $\{v\}^\perp$ in lemma 2.4.9 then follows in both cases directly from lemma 2.4.3. For all bases that we use, such as the nodal Lagrange basis on tetrahedral meshes, the condition on the basis expansion of v in lemma 2.4.9 is then also easily checked. This shows that lemma 2.4.9 can be applied.

In this thesis, the resulting linear equation with a positive definite matrix is then always solved using a preconditioned conjugate gradient solver with an algebraic multigrid preconditioner [173, 23].

2.4.2 The TMS forward problem

We now want to derive a weak formulation for the scalar potential u defining the secondary electric field in the TMS forward problem from definition 2.2.3. If φ denotes a test function, we can proceed in the standard way via

$$\begin{aligned} \int_{\Omega} -f'(t) \cdot \operatorname{div}(\sigma A^S) \cdot \varphi \, dV &= \int_{\Omega} \operatorname{div}(\sigma \nabla u) \varphi \, dV = \int_{\partial\Omega} \langle \sigma \nabla u, \eta \rangle \varphi \, dS - \int_{\Omega} \langle \sigma \nabla u, \nabla \varphi \rangle \, dV \\ &= \int_{\partial\Omega} -f'(t) \langle \sigma A^S, \eta \rangle \varphi \, dS - \int_{\Omega} \langle \sigma \nabla u, \nabla \varphi \rangle \, dV, \end{aligned}$$

which, after reordering, gives

$$\int_{\Omega} \langle \sigma \nabla u, \nabla \varphi \rangle \, dV = -f'(t) \cdot \left(\int_{\partial\Omega} \langle \sigma A^S, \eta \rangle \varphi \, dS - \int_{\Omega} \operatorname{div}(\sigma A^S) \cdot \varphi \, dV \right).$$

Applying partial integration again yields

$$\int_{\partial\Omega} \langle \sigma A^S, \eta \rangle \varphi \, dS - \int_{\Omega} \operatorname{div}(\sigma A^S) \cdot \varphi \, dV = \int_{\Omega} \langle \sigma A^S, \nabla \varphi \rangle \, dV,$$

which finishes the derivation.

Definition 2.4.10 (Weak TMS forward problem). Assume we perform TMS stimulation with a coil generating a magnetic vector potential $A(x, t) = f(t) \cdot A^S(x)$. Then, the generated electric field is described by

$$E = -f'(t) \cdot A^S(x) - \nabla u,$$

where u is given by the problem

$$\text{Find } u \in H^1(\Omega) \text{ such that } \int_{\Omega} \langle \sigma \nabla u, \nabla v \rangle dV = -f'(t) \cdot \int_{\Omega} \langle \sigma A^S, \nabla v \rangle dV \text{ for all } v \in H^1(\Omega).$$

Remark 2.4.11. Note that one could alternatively write the right-hand side in the above definition as

$$-f'(t) \cdot \int_{\Omega} \langle \sigma A^S, \nabla v \rangle dV = \int_{\Omega} \langle \sigma \left(-\frac{\partial A}{\partial t} \right), \nabla v \rangle dV = \int_{\Omega} \langle \sigma E^P, \nabla v \rangle dV.$$

Furthermore, note that if $u^S \in H^1(\Omega)$ solves

$$\int_{\Omega} \langle \sigma \nabla u^S, \nabla v \rangle dV = \int_{\Omega} \langle \sigma A^S, \nabla v \rangle dV \text{ for all } v \in H^1(\Omega),$$

then we have $u(x, t) = -f'(t) \cdot u^S(x)$, similar to the decomposition of A , which in turn also implies $E(x, t) = -f'(t) \cdot (A^S(x) + \nabla u^S(x))$. In particular, if we want to simulate the time course of the electric field generated by TMS, we only need a single FEM run.

Theorem 2.4.12. Assume we perform TMS stimulation with a coil generating a magnetic vector potential $A(x, t) = f(t) \cdot A^S(x)$. Furthermore, assume that $A^S \in L^2(\Omega)^3$. Then, for each timepoint t the TMS weak formulation 2.4.10 has a solution $u_0 \in H^1(\Omega)$. This solution is unique up to an additive constant, i.e., the set of all solutions is given by $u_0 + \mathbb{R} \cdot 1$.

Proof. First, define

$$l : H^1(\Omega) \rightarrow \mathbb{R}; v \mapsto -f'(t) \cdot \int_{\Omega} \langle \sigma A^S, \nabla v \rangle dV.$$

Since the components of A^S are square integrable, l is bounded in $H^1(\Omega)$. Furthermore, since $\nabla 1 = 0$ we also have $l(1) = 0$. But then, the statement of the theorem can be derived by using a strategy completely analogous to the proof of theorem 2.4.5, i.e. reducing to $H_*^1(\Omega)$, where the bilinear form is bounded and coercive by lemma 2.4.3, applying Lax-Milgram there, and showing that this produces a solution in $H^1(\Omega)$ that is unique up to an additive constant. \square

A magnetic dipole at position $z \notin \Omega$ with moment n generates a vector potential

$$A^S(y) = \frac{\mu_0}{4\pi} \cdot n \times \frac{y - z}{\|y - z\|^3}, \quad (2.17)$$

see e.g. [141]. Thus, the vector potential of a magnetic dipole is smooth, and in particular square integrable, on Ω . It follows that the common approach of modelling a coil by a linear combination of magnetic dipoles, as e.g. presented in [47], is covered by theorem 2.4.12.

Furthermore, note that using the same argument as for the PEM tDCS approach, we can apply lemma 2.4.9 when performing a TMS finite element simulation.

2.5 Deriving FEM formulations for the measurement forward problems

We thus finally arrive at the problem of assigning finite element formulations to the EEG and MEG forward problems, see definitions 2.1.1, 2.1.2, and 2.1.3. The first question, whose answer will in fact be one of the major sources of difficulty in what follows, is how to model the neural activity j^P .

The current standard in the MEEG community is to model the neural activity j^P as a linear combination of so-called *current dipoles*, see e.g. [42]. A current dipole encodes a current at a certain position with a certain direction and strength. Concretely, a current dipole is characterized by a *dipole position*, say $x_0 \in \Omega$, and a *dipole moment*, say $M \in \mathbb{R}^3$, and is given by the distribution $M \cdot \delta_{x_0}$, where δ_{x_0} denotes the Dirac distribution at position x_0 . It has been shown that current dipoles can adequately describe the generators of human brain activity [40, 41]. We want to note that there have been investigations into replacing the current dipole with a more sophisticated approach, such as using a multipolar source model [101, 19]. These approaches have, however, to the best of our knowledge, not gained any real traction in the community. Accordingly, we will for the rest of this thesis assume a dipolar source model.

Now, a close look at the definitions of the measurement forward problems shows that they depend linearly on the neural activity j^P . As we model the neural activity as a linear combination of current dipoles, we see that the EEG and MEG forward problems reduce to solving for the electric potential and magnetic fields generated by a single point dipole.

Looking at definition 2.1.3, we see that to solve the MEG forward problem, we first need to solve the EEG forward problem. We will thus first focus on the EEG forward problem, i.e., the computation of electric potentials in the head due to a current dipole.

There are, roughly speaking, two classes of approaches for the EEG forward problem, the so-called *direct approaches* and the so-called *subtraction approaches*. In the direct approaches, one poses a forward problem whose solution “directly” is an approximation of the electric potential u . In the following, we will see that, due to the singularity of the dipole, this is typically done by substituting the dipole, or its divergence, with a more regular object. In subtraction approaches, on the other hand, one does not try to directly approximate the potential u , but instead tries to approximate $u - \varphi$ for a suitable φ . These types of approaches have been shown to both possess a rigorous mathematical foundation and produce accurate simulation results [239, 53, 18]. Nevertheless, they were rarely used in practice due to their high computational costs. In this thesis, we will present a further development of this category of approaches, which will lead to the so-called *local subtraction approach*. This approach reduces the computational costs of subtraction approaches to an extent that they become usable in practice. This is demonstrated in chapter 4 of this thesis, where the source reconstructions for an epilepsy patient are based on EEG and MEG forward solutions computed using the local subtraction approach. This approach was recently published in [91]. Our discussion here will largely follow the one given in the publication, with the exception that in [91] we only discuss the PEM case, while here we will also discuss the CEM case. The extension of the local subtraction approach to the CEM case was developed by us in close cooperation with Fatih Karpuz during the course of his master’s thesis [103].

We will now first introduce the direct EEG approaches, as they are the current state-of-the-art, and our new approach thus needs to be validated against them. Then, we will derive the local subtraction EEG approach and discuss how it relates to the previous iterations of subtraction-type approaches. Finally, we will discuss the MEG forward problem.

2.5.1 Direct approaches for the EEG forward problem

We start in the standard way. Looking at definitions 2.1.1 and 2.1.2, we get in both the PEM and the CEM case for a test function φ that

$$\int_{\Omega} \operatorname{div}(j^P) \varphi \, dV = \int_{\Omega} \operatorname{div}(\sigma \nabla u) \varphi \, dV = \int_{\partial\Omega} \langle \sigma \nabla u, \eta \rangle \varphi \, dS - \int_{\Omega} \langle \sigma \nabla u, \nabla \varphi \rangle \, dV. \quad (2.18)$$

In the PEM case, this yields

$$\int_{\Omega} \langle \sigma \nabla u, \nabla \varphi \rangle \, dV = - \int_{\Omega} \operatorname{div}(j^P) \varphi \, dV. \quad (2.19)$$

Looking at definition 2.1.2, we have in the CEM case for an arbitrary vector $c = (c_1, \dots, c_L) \in \mathbb{R}^L$ that

$$\begin{aligned} \int_{\partial\Omega} \langle \sigma \nabla u, \eta \rangle \varphi \, dS &= \sum_{l=1}^L \int_{e_l} \langle \sigma \nabla u, \eta \rangle \varphi \, dS = \sum_{l=1}^L \int_{e_l} \langle \sigma \nabla u, \eta \rangle (\varphi - c_l) \, dS \\ &= \sum_{l=1}^L \frac{1}{z_l} \int_{e_l} (U_l - u) (\varphi - c_l) \, dS, \end{aligned}$$

which together with (2.18) yields

$$\int_{\Omega} \langle \sigma \nabla u, \nabla \varphi \rangle \, dV + \sum_{l=1}^L \frac{1}{z_l} \int_{e_l} (u - U_l) (\varphi - c_l) \, dS = - \int_{\Omega} \operatorname{div}(j^P) \varphi \, dV.$$

If j^P were an element of $H(\operatorname{div}; \Omega)$, we would now be done with the derivation. But since we model the neural activity based on current dipoles, we have an additional layer of complexity.

The partial integration approach

The naive approach would be to just use the dipole $j^P = M \cdot \delta_{x_0}$ as is. By the definition of derivatives of distributions, we have for a sufficiently regular test function φ that

$$\begin{aligned} -\operatorname{div}(M \cdot \delta_{x_0})(\varphi) &= \sum_{i=1}^3 -\frac{\partial}{\partial x_i} (M_i \cdot \delta_{x_0})(\varphi) = \sum_{i=1}^3 M_i \cdot \delta_{x_0} \left(\frac{\partial \varphi}{\partial x_i} \right) = \sum_{i=1}^3 M_i \cdot \frac{\partial \varphi}{\partial x_i}(x_0) \\ &= \langle M, \nabla \varphi(x_0) \rangle. \end{aligned}$$

In the community, this is known as the *partial integration approach*.

Definition 2.5.1 (Partial integration approach). Assume we have a current dipole at position $x_0 \in \Omega$ with moment $M \in \mathbb{R}^3$. Let $V_h \subset H^1(\Omega)$ be a finite-dimensional subspace such that for each $u \in V_h$ we have that u is C^1 in an environment of x_0 . Let a_{PEM} and a_{CEM} be defined as in (2.12) and (2.13). Then, the PEM partial integration approach is given by

$$\text{Find } u_h \in V_h \text{ such that } a_{\text{PEM}}(u_h, v_h) = \langle M, \nabla v_h(x_0) \rangle \text{ for all } v_h \in V_h.$$

Furthermore, the CEM partial integration approach is given by finding $(u_h, U_h) \in V_h \oplus \mathbb{R}^L$ such that

$$a_{\text{CEM}}((u_h, U_h), (v_h, V_h)) = \langle M, \nabla v_h(x_0) \rangle$$

for all $(v_h, V_h) \in V_h \oplus \mathbb{R}^L$.

Note that, since $\langle M, \nabla 1 \rangle = 0$, we can again apply lemma 2.4.9 to solve the FEM problem.

It turns out that this approach produces relatively large errors, see e.g. [116]. The reason for this is that the highly singular behaviour of the focal current dipole cannot be accurately captured by typical finite element trial spaces. As a concrete example of this, note that, e.g., for piecewise linear trial functions (i.e., when using a nodal Lagrange basis on a tetrahedral mesh), the right-hand side functional defined by 2.5.1 does not change when moving the dipole position within the same element.

Different approaches have been suggested to remedy this. These typically work by substituting the dipole with a regularized object that behaves similarly to a dipole, but spreads the behaviour of the dipole over a larger portion of the head domain. The most common strategy to do this is the so-called *Venant approach*.

The Venant approach

The general idea is relatively simple. We decide on a set of “trial” functions $\{\varphi_1, \dots, \varphi_N\}$, and want to find coefficients $c_1, \dots, c_N \in \mathbb{R}$ such that

$$-\operatorname{div}(M \cdot \delta_{x_0}) \approx \sum_{j=1}^N c_j \cdot \varphi_j. \quad (2.20)$$

We can then use $\sum_{j=1}^N c_j \cdot \varphi_j$ instead of $-\operatorname{div}(M \cdot \delta_{x_0})$ in the FEM simulation, and, as the union of the supports of $\varphi_1, \dots, \varphi_N$ typically extends over multiple mesh elements, this spreads out the dipole singularity over a wider part of the head model. As described above, this tends to reduce the error of the simulation.⁶ All published variants of the Venant approach choose the trial functions by selecting a set of positions $x_1, \dots, x_N \in \Omega$ and setting $\varphi_j := \delta_{x_j}$, where δ_{x_j} denotes the Dirac distribution at x_j . In the Venant approach, one thus approximates the divergence of the current dipole by a linear combination of monopoles.

To choose c_1, \dots, c_N , we select a set of “test” functions $T = \{\psi_1, \dots, \psi_M\}$, and determine c_1, \dots, c_N based on the conditions

$$\int_{\Omega} -\operatorname{div}(M \cdot \delta_{x_0}) \cdot \psi_i \, dV = \int_{\Omega} \left(\sum_{j=1}^N c_j \cdot \varphi_j \right) \cdot \psi_i \, dV \quad (2.21)$$

for $1 \leq i \leq M$. All published variants of the Venant approach select the set of test functions as a set of linear combinations of elements from $\{(x - x_0)^\alpha \mid \alpha \in \mathbb{N}_0^3\}$. The original version of the Venant approach from [28] corresponds to the choice

$$T_B := \{1, (x - x_0)^{e_1}, (x - x_0)^{e_2}, (x - x_0)^{e_3}, (x - x_0)^{2 \cdot e_1}, (x - x_0)^{2 \cdot e_2}, (x - x_0)^{2 \cdot e_3}\}, \quad (2.22)$$

where the subscript “B” is supposed to stand for “Buchner”. In the dissertation [143], it is instead suggested to use

$$T_N := \{(x - x_0)^\alpha \mid \|\alpha\|_\infty \leq 2\}. \quad (2.23)$$

We want to note that this version of the Venant approach is the default choice in the pipeline that we published to Zenodo [54]. The “N” is supposed to stand for “Nüßing”, named after the author

⁶We want to emphasize that the “spreading out” is really the central reason why Venant approaches generally reduce errors. If the supports of the trial functions are all contained in the dipole element, one can show that all the different versions of the Venant approach that are presented in this chapter are equivalent to the partial integration approach. A corresponding proof can be found in [143], Corollar 2.1.

of the aforementioned PhD thesis. Another choice is given by the *multipolar Venant* approach, which corresponds to the choice

$$T_M := \{1, (x - x_0)^{e_1}, (x - x_0)^{e_2}, (x - x_0)^{e_3}, 3(x - x_0)^{e_1+e_2}, 3(x - x_0)^{e_1+e_3}, 3(x - x_0)^{e_2+e_3}, \\ 3(x - x_0)^{2 \cdot e_1} - \|x - x_0\|_2^2, 3(x - x_0)^{2 \cdot e_2} - \|x - x_0\|_2^2, 3(x - x_0)^{2 \cdot e_3} - \|x - x_0\|_2^2\}. \quad (2.24)$$

This approach was first introduced in [86] and published in [218]. While T_B and T_N are, to the best of our knowledge, ad hoc choices, T_M is based on a Physics-inspired heuristic, specifically the so-called *multipole expansion*, see e.g. [141]. Concretely, we have already previously used the fact that for a sufficiently regular $\rho : \mathbb{R}^3 \rightarrow \mathbb{R}$ the unique solution f of $\Delta f = -\rho$ decaying at infinity is given by

$$f(x) = \frac{1}{4\pi} \int_{\mathbb{R}^3} \frac{\rho(y)}{\|x - y\|} dy, \quad (2.25)$$

see e.g. [64], §16, Satz 1. But now we have

$$\frac{1}{\|x - y\|} = \frac{1}{(\|x\|^2 - 2 \cdot \langle x, y \rangle + \|y\|^2)^{\frac{1}{2}}} = \frac{1}{\|x\|} \cdot \left(1 - \frac{2 \langle x, y \rangle}{\|x\|^2} + \frac{\|y\|^2}{\|x\|^2}\right)^{-\frac{1}{2}}$$

If we now assume that ρ has support close to $0 \in \mathbb{R}^3$, and that we want to evaluate f at a position x with a large distance from zero, we have $\frac{\|y\|}{\|x\|} \approx 0$ for y in the support of ρ , and then also $-\frac{2 \langle x, y \rangle}{\|x\|^2} + \frac{\|y\|^2}{\|x\|^2} \approx 0$. Then, applying a standard Taylor expansion to the analytic function $h(t) = (1 + t)^{-\frac{1}{2}}$ yields that

$$h(t) = \sum_{n=0}^{\infty} \frac{h^{(n)}(0)}{n!} t^n = \frac{h(0)}{0!} + \frac{h'(0)}{1!} t + \frac{h''(0)}{2!} t^2 + \mathcal{O}(t^3) \\ = 1 - \frac{1}{2} t + \frac{3}{8} t^2 + \mathcal{O}(t^3).$$

Using this expansion for $t = -\frac{2 \langle x, y \rangle}{\|x\|^2} + \frac{\|y\|^2}{\|x\|^2}$, gathering terms based on their decay rate, and inserting into (2.25), we get

$$4\pi \cdot f(x) = \int_{\mathbb{R}^3} \rho(y) dy \cdot \frac{1}{\|x\|} + \frac{\langle \int_{\mathbb{R}^3} y \cdot \rho(y) dy, x \rangle}{\|x\|^3} + \frac{1}{2} \frac{\langle x, \int_{\mathbb{R}^3} \rho(y) \cdot (3yy^\top - \|y\|^2 \cdot E_3) dy \cdot x \rangle}{\|x\|^5} \\ + \mathcal{O}\left(\frac{1}{\|x\|^4}\right).$$

One typically calls

$$q := \int_{\mathbb{R}^3} \rho(y) dy$$

the *monopole moment* of ρ ,

$$M := \int_{\mathbb{R}^3} y \cdot \rho(y) dy$$

the *dipole moment* of ρ , and

$$Q := \int_{\mathbb{R}^3} \rho(y) \cdot (3yy^\top - \|y\|^2 \cdot E_3) dy$$

the *quadrupole moment* of ρ . Thus, if we want to approximate ρ by ν in such a way that the solutions f, g of $\Delta f = -\rho$ and $\Delta g = -\nu$ differ only up to an order of $\mathcal{O}(\frac{1}{\|x\|^4})$ when looking at a large distance from their supports, this can be achieved by choosing ν in such a way that it matches the monopole, dipole, and quadrupole moments of ρ . Looking at (2.24), we see that the multipolar Venant approach does exactly this.

Furthermore, summing the last three test functions in T_M yields

$$\begin{aligned} (3(x - x_0)^{2 \cdot e_1} - \|x - x_0\|_2^2) + (3(x - x_0)^{2 \cdot e_2} - \|x - x_0\|_2^2) + (3(x - x_0)^{2 \cdot e_3} - \|x - x_0\|_2^2) \\ = 3 \cdot \|x - x_0\|^2 - 3 \cdot \|x - x_0\|^2 = 0, \end{aligned}$$

i.e. the functions in T_M are linearly dependent. Thus, choosing T_M yields only 9 independent conditions, and not 10, as e.g. claimed in [218]. To the best of our knowledge, this is currently not known in the EEG forward modeling community.

This finishes the description of the typical choices for the set of trial functions $\{\varphi_1, \dots, \varphi_N\}$ and the set of test functions $\{\psi_1, \dots, \psi_M\}$ for the Venant approach. We now continue with implementation aspects of the Venant approach. To this end, we first compute how a current dipole acts against the specified set of possible test functions.

Lemma 2.5.2. *Let $\alpha \in \mathbb{N}^3$. We then have*

$$-\operatorname{div}(M \cdot \delta_{x_0})((x - x_0)^\alpha) = \begin{cases} 0 & \text{if } |\alpha|_1 \neq 1 \\ M_i & \text{if } \alpha = e_i \text{ for } 1 \leq i \leq 3 \end{cases}$$

Proof. If $\alpha = 0$, we have $-\operatorname{div}(M \cdot \delta_{x_0})(1) = \langle M, \nabla 1 \rangle = 0$. For $1 \leq i \leq 3$, we have

$$-\operatorname{div}(M \cdot \delta_{x_0})((x_i - (x_0)_i)) = \langle M, \nabla(x_i - (x_0)_i)(x_0) \rangle = M_i.$$

If $|\alpha|_1 > 1$, at least one of the following two statements must be true.

- There are $i \neq j$ with $\alpha_i \geq 1$ and $\alpha_j \geq 1$.
- There is i such that $\alpha_i \geq 2$.

In the second case, it follows that for each $1 \leq k \leq 3$ that $\frac{\partial(x-x_0)^\alpha}{\partial x_k}$ can be factorized as $(x - x_0)^{e_i} \cdot f(x)$ for some polynomial f . But this implies $\frac{\partial(x-x_0)^\alpha}{\partial x_k}(x_0) = 0$, and thus we have $\nabla(x - x_0)^\alpha(x_0) = 0$. In the first case, the partial derivatives for $k \neq i$ still factor through $(x - x_0)^{e_i}$, while the partial derivative with respect to x_i factors through $(x - x_0)^{e_j}$, and we arrive at the same conclusion as before. \square

Now, let $T = \{\psi_1, \dots, \psi_M\}$ denote our choice of test function set, i.e. either $T = T_B$, $T = T_N$, or $T = T_M$. Lemma 2.5.2 can be used for the computation of the vector $b = (b_i)_{1 \leq i \leq M} \in \mathbb{R}^M$ defined by $b_i = -\operatorname{div}(M \cdot \delta_{x_0})(\psi_i)$. Furthermore, if x_1, \dots, x_N denote the positions defining the trial functions, we have

$$\delta_{x_i}((x - x_0)^\alpha) = (x_i - x_0)^\alpha.$$

If we now define $A = (a_{ij}) \in \mathbb{R}^{M \times N}$ by $a_{ij} = \delta_{x_j}(\psi_i)$, we see that (2.21) is equivalent to $A \cdot c = b$, where $c = (c_1, \dots, c_N) \in \mathbb{R}^N$ is the vector containing the interpolation coefficients.

One generally chooses the positions $x_1, \dots, x_N \in \Omega$ by selecting element nodes in the vicinity of the dipole position. Depending on the local mesh structure and on the choice of test functions,

both $N \leq M$ as well as $N > M$ are possible. To give each case a unified treatment, one relaxes the condition $A \cdot c = b$ from searching for an exact solution to searching for a Tikhonov regularized least squares solution. Furthermore, since matching the lower order moments, i.e. the monopole moment and the dipole moment, is more important for an accurate simulation than matching the higher order moments, one introduces an additional weighting by specifying a constant $s > 0$.⁷ Then, if the i -th test function ψ_i is a polynomial of degree k_i , one scales the corresponding equation by s^{-k_i} . Concretely, setting $S = \text{diag}(s^{-k_1}, \dots, s^{-k_M}) \in \mathbb{R}^{M \times M}$, one defines $\bar{A} := S \cdot A$ and $\bar{b} := S \cdot b$, and searches for regularized least squares solutions of $\bar{A} \cdot c = \bar{b}$.

In the Tikhonov penalty term, we want to penalize large monopolar loads at positions with a great distance from the position of the current dipole. We thus set

$$\bar{W} = \frac{1}{s} \cdot \text{diag}(\|x_1 - x_0\|^{\frac{r}{2}}, \dots, \|x_N - x_0\|^{\frac{r}{2}}) \in \mathbb{R}^{N \times N}$$

where $r > 0$ is a user-controlled parameter. Putting all of this together, we define the Venant weights $c_V \in \mathbb{R}^N$ by the condition

$$c_V := \arg \min_{c \in \mathbb{R}^N} \|\bar{A} \cdot c - \bar{b}\|_2^2 + \lambda \cdot \|\bar{W} \cdot c\|_2^2, \quad (2.26)$$

where $\lambda > 0$ is a regularization parameter. By standard theory (see e.g. [203], Chapter 2), problem (2.26) has a unique solution that can be computed via

$$c_V = \left(\bar{A}^\top \cdot \bar{A} + \lambda \cdot \bar{W}^\top \bar{W} \right)^{-1} \cdot \bar{A}^\top \cdot \bar{b}.$$

This finishes the derivation of the Venant approach.

Definition 2.5.3 (Venant approach). We use the notation defined in the preceding paragraphs. Let $V_h \subset H^1(\Omega)$ be a finite-dimensional subspace such that for each $u \in V_h$ we have that u is continuous in an environment of the dipole position x_0 . Let a_{PEM} and a_{CEM} be defined as in (2.12) and (2.13). Then, the PEM Venant approach is given by

$$\text{Find } u_h \in V_h \text{ such that } a_{\text{PEM}}(u_h, v_h) = \sum_{j=1}^N (c_V)_j \cdot v_h(x_j) \text{ for all } v_h \in V_h.$$

Furthermore, the CEM partial integration approach is given by finding $(u_h, U_h) \in V_h \oplus \mathbb{R}^L$ such that

$$a_{\text{CEM}}((u_h, U_h), (v_h, V_h)) = \sum_{j=1}^N (c_V)_j \cdot v_h(x_j)$$

for all $(v_h, V_h) \in V_h \oplus \mathbb{R}^L$.

In principle, we would now like to again apply lemma 2.4.9 to show the existence of a solution for the Venant approach. The critical condition here is $l(1) = 0$ for the right-hand side. In lemma 2.4.9, it is shown that this is a sufficient condition for the existence of a solution. But it is in fact also necessary, as $a(u, v) = l(v)$ for all test functions v implies that

$$l(1) = a(u, 1) = 0,$$

⁷In the corresponding publications [28, 218], and also in the implementation in our DUNEuro toolbox, this parameter is typically called either *reference length* or *reference distance*.

both for $a = a_{\text{PEM}}$ as well as for $a = a_{\text{CEM}}$. We thus see that if $l(1) \neq 0$, there is no solution of the problem $a(u, v) = l(v)$ for all test functions v .

Looking at definition 2.5.3, we see that the existence of a solution for the Venant approach is equivalent to $\sum_{j=1}^N (c_V)_j = 0$. Now, the coefficients c_V are defined by (2.26). If we choose a very small λ , this essentially comes down to minimizing $\|\bar{A} \cdot c - \bar{b}\|_2^2$. Furthermore, each basis of Venant test functions contains the constant 1 function, whose corresponding row in this problem corresponds to $\left(\sum_{j=1}^N (c_V)_j\right)^2$. Additionally, the remaining rows contain the scaling factor s^{-k_i} , where k_i is the degree of the test function, and hence (for the common choice of $s = 20$ mm) get comparatively less weight than the row corresponding to the constant 1 function. We thus see that the minimization problem defining the Venant coefficients c_V places a high weight on keeping $\left(\sum_{j=1}^N (c_V)_j\right)^2$ small. Nevertheless, even if $M < N$ and the equation $\bar{A} \cdot c = \bar{b}$ can be solved exactly, the presence of the Tikhonov regularization leads to deviations of $\bar{A} \cdot c$ from \bar{b} . Running the actual computations, one can easily check that $\left(\sum_{j=1}^N (c_V)_j\right)^2$ generally is small, but not zero, and is in particular much larger than what can arise due to rounding errors. This deviation gets larger with larger values for λ . We have thus arrived at the following.

The Venant approach as defined in 2.5.3 in general does not have a solution.

We have, however, shown in lemma 2.4.9 that the modified stiffness matrix (\hat{A} in that lemma) is positive definite. In particular, we can just apply the conjugate gradient algorithm to the right-hand side defined by 2.5.3, and it will converge to something. Thus, the general response to the non-existence of a solution to the Venant problem is to just ignore it, apply lemma 2.4.9 anyway, and hope that this produces a sensible result. In fact, empirical experience shows that it does, as demonstrated in the corresponding publications [28, 218, 86, 143].

We want to give a more rigorous discussion. The equation $l(1) = 0$ is commonly called the *compatibility condition*. The next lemma describes the consequences of violating the compatibility condition.

Lemma 2.5.4. *We use the definition of \hat{A} and the notation from lemma 2.4.9. Furthermore, let $l : V \rightarrow \mathbb{R}$ denote a linear functional, where we now explicitly do **not** assume $l(v) = 0$. Then, let y and \hat{y} be defined as in lemma 2.4.9, and let $x \in \mathbb{R}^N$ be the solution of $\hat{A} \cdot x = \hat{y}$. Additionally, for the basis $\mathfrak{B} = \{b_1, \dots, b_N\}$ let $\mathfrak{B}^* = \{b_1^*, \dots, b_N^*\}$ denote the corresponding dual basis of V^* , where b_i^* is the unique linear functional with the property $b_i^*(b_j) = \delta_{ij}$, where δ_{ij} denotes the Kronecker delta. Then, x is the coefficient vector of a solution $u \in V$ of*

$$\text{Find } u \in V \text{ such that } a(u, w) = \tilde{l}(w) \text{ for all } w \in V, \quad (2.27)$$

where

$$\tilde{l} := l - \frac{l(v)}{b_1^*(v)} \cdot b_1^* \in V^*.$$

Proof. First, note that

$$\tilde{l}(v) = l(v) - \frac{l(v)}{b_1^*(v)} \cdot b_1^*(v) = 0.$$

Now define $\tilde{y} := (\tilde{l}(b_i))_{1 \leq i \leq N} \in \mathbb{R}^N$, and $\hat{\tilde{y}} := (0, \tilde{y}_2, \dots, \tilde{y}_N) \in \mathbb{R}^N$. By lemma 2.4.9 it then follows that (2.27) has a unique solution whose coefficients $\tilde{x} \in \mathbb{R}^N$ are given by $\hat{A} \cdot \tilde{x} = \hat{\tilde{y}}$ and

that fulfill $\tilde{x}_1 = 0$. But then $\hat{\tilde{y}}_1 = 0 = \hat{y}_1$, and for $2 \leq i \leq N$ we have

$$\hat{\tilde{y}}_i = \tilde{l}(b_i) \stackrel{b_1^*(b_i) = 0 \text{ for } i \geq 2}{=} l(b_i) = \hat{y}_i.$$

But this yields $\hat{\tilde{y}} = \hat{y}$, which shows $\tilde{x} = x$. \square

We want to explicitly state what lemma 2.5.4 means for typical FEM approximations.

Corollary 2.5.5. *Assume we apply the strategy from lemma 2.4.9 for $a = a_{PEM}$ and V the space of piecewise linear functions on a tetrahedral mesh. Then, we typically choose $v = 1$ and \mathfrak{B} as the set of nodal Lagrange basis functions. In particular, we have a basis function for each mesh node, and if n_1, \dots, n_N denote the mesh vertices corresponding to the nodal basis functions, then the dual basis is given by $b_i^* = \delta_{n_i}$, i.e. b_i^* is given by evaluating a trial function at the vertex n_i . In particular, $b_1^*(1) = 1$, and thus the vector x computed by applying the strategy from lemma 2.4.9 describes the solution of the problem*

$$\text{Find } u \in V \text{ such that } a_{PEM}(u, w) = l(w) - l(1) \cdot \delta_{n_1}(w) \text{ for all } w \in V.$$

Speaking less formally, not fulfilling the compatibility condition can, in this case, be interpreted as a violation of charge conservation. In the FEM simulation, the unaccounted charges are artificially compensated by the introduction of a sink of matching strength at an arbitrary mesh node. This is illustrated in figure 2.4.

Furthermore, there is nothing special about either tetrahedral meshes or piecewise affine trial functions in corollary 2.5.5. One e.g. gets essentially the same result for piecewise multilinear trial functions on meshes consisting of cubes.

We have thus derived the Venant approach and discussed its solvability. Before continuing, we want to address two remaining questions.

- Each choice of test basis functions leads to a different right-hand side. We have introduced the selections T_B , T_N , and T_M . This, of course, raises the question of which of these selections one should use. In particular, as there are, to the best of our knowledge, no published comparisons between the Nüßing Venant, i.e. T_N , and the multipolar Venant, i.e. T_M . In section 2.8.3 of this thesis, we will investigate this question. To preemptively give an answer, we will see that, at least for tetrahedral meshes with piecewise affine trial functions, the multipolar Venant, i.e., T_N , generally performs best.
- To actually compute the Venant interpolation vector c_V , one needs to specify a number of parameters. Concretely, one needs to set the weighting factor $s > 0$, the exponent $r > 0$, and the regularization parameter $\lambda > 0$. In this thesis, we follow the recommendation from [218], and choose $s = 20$, $r = 2$, and $\lambda = 10^{-6}$. For s , we assume that the mesh positions are given in millimeters. If the mesh is given in other units, s should be scaled accordingly.

The Whitney approach

As a final class of direct approaches, we want to introduce the class of so-called *Whitney approaches*, which are sometimes also called *$H(\text{div})$ approaches*. This class of approaches was developed over a number of successive papers [161, 159, 15, 162, 133].

The general idea is quite similar to the Venant approach. While in the Venant approach, one chooses a set of *scalar* “trial” functions $\{\varphi_1, \dots, \varphi_N\}$ and tries to select coefficients c_1, \dots, c_N such

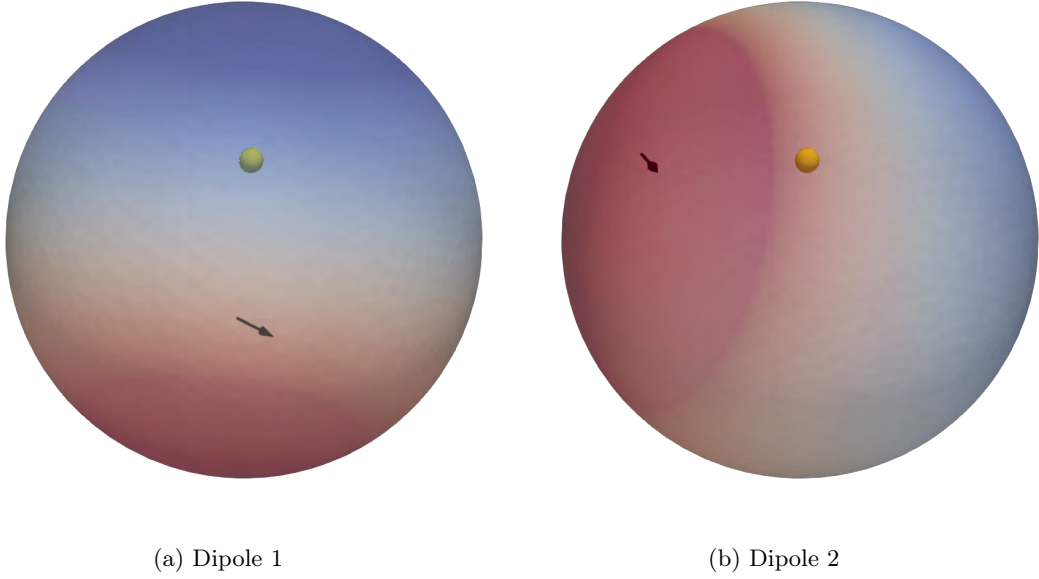


Figure 2.4: Illustration of a gross violation of the compatibility condition for the EEG forward problem. To generate these visualizations, we took a spherical model and two dipoles, shown as black arrows in the figures above. For each of these dipoles, we then computed the vector $c_V \in \mathbb{R}^N$ for the multipolar Venant test function basis selection. But before solving the corresponding linear system, we added a large constant to $(c_V)_1$, making it approximately $2 \cdot 10^3$ times larger than the remaining entries of c_V . The resulting functional $f(v) = \sum_{i=1}^N (c_V)_i \cdot v(x_i)$ was then used with the strategy described in lemma 2.4.9. Note that this strategy is the default behaviour of DUNEuro. The resulting FEM solution was then plotted over the spherical model. Finally, the mesh node that was assigned the index 1 is visualized as a yellow sphere. We see that the results behave as one would expect from corollary 2.5.5, i.e., the result essentially shows a dipolar pattern, corresponding to two monopolar loads, one placed at the true dipole position, and one at the position of the vertex with index 1. Note that the purpose of this figure is to present an understandable demonstration of corollary 2.5.5, and is hence based on extreme values. In actual Venant simulations, the effect is much smaller and is, in fact, not really noticeable. Nevertheless, in the Venant approach, a deviation of $\sum_{i=1}^N (c_V)_i$ from 0 is compensated by placing a sink with the strength of this deviation at an unrelated position in the mesh.

that

$$-\operatorname{div}(M \cdot \delta_{x_0}) \approx \sum_{j=1}^N c_j \varphi_j,$$

the Whitney approach instead chooses a set of *vector* “trial” functions $\{\psi_1, \dots, \psi_N\} \in H(\operatorname{div}; \Omega)$ and tries to select c_1, \dots, c_N such that

$$M \cdot \delta_{x_0} \approx \sum_{j=1}^N c_j \psi_j.$$

Since by construction the vector trial functions are contained in $H(\text{div}; \Omega)$, the divergence $-\text{div}(\sum_{j=1}^N c_j \psi_j)$ is a square integrable function, and can then be used as a more regular replacement of $-\text{div}(M \cdot \delta_{x_0})$ in the EEG forward problem.

Definition 2.5.6 (Whitney approach). Let $V_h \subset H^1(\Omega)$ denote a finite-dimensional subspace. Assume we have a dipole at position $x_0 \in \Omega$ with moment $M \in \mathbb{R}^3$. Furthermore, let $\{\psi_1, \dots, \psi_N\} \subset H(\text{div}; \Omega)$ and $c_1, \dots, c_N \in \mathbb{R}^N$ be given such that $M \cdot \delta_{x_0} \approx \sum_{j=1}^N c_j \psi_j$. We assume that the supports of ψ_1, \dots, ψ_N are contained in the interior of Ω . Let a_{PEM} and a_{CEM} be defined as in (2.12) and (2.13). Then, the PEM Whitney approach is given by the problem of finding $u_h \in V_h$ such that

$$a_{\text{PEM}}(u_h, v_h) = - \int_{\Omega} \text{div} \left(\sum_{j=1}^N c_j \cdot \psi_j \right) \cdot v_h \, dV$$

for all $v_h \in V_h$. The CEM Whitney approach is accordingly given by the problem of finding $(u_h, U_h) \in V_h \oplus \mathbb{R}^L$ such that

$$a_{\text{CEM}}((u_h, U_h), (v_h, V_h)) = - \int_{\Omega} \text{div} \left(\sum_{j=1}^N c_j \cdot \psi_j \right) \cdot v_h \, dV$$

for all $(v_h, V_h) \in V_h \oplus \mathbb{R}^L$.

Since by assumption the supports of ψ_1, \dots, ψ_N are contained in the interior of Ω , we have

$$- \int_{\Omega} \text{div} \left(\sum_{j=1}^N c_j \cdot \psi_j \right) \cdot 1 \, dV = - \int_{\partial\Omega} \left\langle \sum_{j=1}^N c_j \cdot \psi_j, \eta \right\rangle \, dS = 0.$$

Thus, the compatibility condition is fulfilled, and we can apply lemma 2.4.9 to solve the FEM problem.

This description leaves open how to choose the vector functions ψ_1, \dots, ψ_N and the coefficients c_1, \dots, c_N .

In the aforementioned papers, the Whitney approach is only defined for tetrahedral meshes, where one employs piecewise linear trial functions for the space V_h . There is no principal reason to do this, as one could also choose vector functions ψ_1, \dots, ψ_N on cube meshes, e.g., similar to what is done in [152]. This might be a topic for future investigations, but for now, we will follow the literature and also restrict ourselves to tetrahedral meshes with piecewise linear trial functions.

The candidates for the vector functions ψ_1, \dots, ψ_N are then, to the best of our knowledge, an ad hoc selection out of a list of $H(\text{div}; \Omega)$ -functions from [2]. Concretely, the Whitney approach utilizes the so-called *edge-based face functions* and *edge-based interior functions* from table 4 in [2]. In [2], these functions depend on a parameter $l \geq 0$, where the Whitney approach takes the functions corresponding to $l = 0$.

We could now give the definitions of these functions and derive their properties. We will instead restrict ourselves to presenting the punchlines and refer to [162] for the computational details. The central insight is that for the purpose of a tetrahedral P1 FEM simulation, both the edge-based face functions as well as the edge-based interior functions are completely determined from just two points. For the edge-based interior functions, these two points are the endpoints of the corresponding edge. For the edge-based face functions, let F be a common face of the two

tetrahedrons T_1 and T_2 . Then, T_1 has exactly one vertex that is not contained in F , and T_2 also has exactly one vertex that is not contained in F . The edge-based face function corresponding to this face F is then determined from these two vertices that are not contained in F .

Lemma 2.5.7. *Assume we have a tetrahedral mesh of the head domain Ω . Let $p_1, p_2 \in \Omega$ denote two mesh vertices defining either an edge-based face function or an edge-based interior function as defined in [162], and call this function ψ . Then, there exists $\lambda \neq 0$ such that if φ is a continuous function that is piecewise linear on each mesh tetrahedron, i.e., a P1 FEM test function, we have that*

$$\begin{aligned} - \int_{\Omega} \operatorname{div}(\psi) \cdot \varphi \, dV &= \lambda \cdot (\delta_{p_2} - \delta_{p_1})(\varphi) \\ &= \lambda \cdot \varphi(p_2) - \lambda \cdot \varphi(p_1). \end{aligned}$$

Proof. [162], appendices B and C. □

We thus see that in the context of a tetrahedral P1 FEM, the $H(\operatorname{div}; \Omega)$ functions of the Whitney approach are exactly equivalent to placing two monopolar loads of opposite strength on two mesh vertices that are connected by an element edge or face. In particular, we see that the Whitney approach as defined in 2.5.6 is equivalent to substituting the divergence of the dipole with a cloud of monopoles placed on mesh vertices in the proximity of the dipole position.

Note that this is essentially the same process as in the Venant approach. As a consequence, one could thus use the same strategies as in the Venant approach to compute the interpolation coefficients c_1, \dots, c_N . The authors of the papers developing the Whitney approach, however, decided to develop their own strategies to perform this interpolation, which we will now discuss.

First of all, if ψ is either an edge-based face function or an edge-based interior function corresponding to the mesh vertices p_1 and p_2 , the authors in [162] show that

$$\int_{\Omega} \psi \, dV = \frac{p_2 - p_1}{\|p_2 - p_1\|},$$

i.e., the mean direction of the vector field ψ points in the direction of the difference of the two mesh vertices. The general idea of the interpolation is now to replace ψ with a dipole at some position x with moment $\frac{p_2 - p_1}{\|p_2 - p_1\|}$, and use this dipole for the computation of the interpolation coefficients c_1, \dots, c_N . We want to emphasize that these dipoles do not directly enter the FEM computation, but are only auxiliary tools for the computation of the interpolation coefficients c_1, \dots, c_N . This, of course, raises the question of how to assign a position x to the vector field ψ . In [161] and [15], the authors investigate this question by running simulations on a set of candidate positions, and conclude that the midpoint $\frac{p_1 + p_2}{2}$ yields the best results.

We thus arrive at the following strategy, as described in [162]. Given a dipole at position $x_0 \in \Omega$ with moment $M \in \mathbb{R}^3$, we first locate the tetrahedral element T containing the position x_0 . Then, we assign to each of its four faces an edge-based face function, and to each of its six edges an edge-based interior function. Then, we iterate over each of the four elements sharing a face with T and assign to each of its edges an edge-based interior function. In the typical case, we thus get four face functions, six edge-based functions from the original tetrahedron, and three additional edge functions for each of the four tetrahedrons sharing a face with T , for a total of 22 functions. Note that this number can be smaller than 22, as two elements that share a face with T might also share an edge with each other.

We thus get vector-valued functions ψ_1, \dots, ψ_N , where each function ψ_i is determined from two positions p_1^i and p_2^i . To the function ψ_i we then assign the dipole at position $x_i = \frac{p_1^i + p_2^i}{2}$ with moment $M_i = \frac{p_2^i - p_1^i}{\|p_2^i - p_1^i\|}$. Based on these dipoles, we can now compute the interpolation coefficients c_1, \dots, c_N . For this, two approaches have been suggested, first the so-called *position based optimization (PBO)* [15], and secondly the so-called *mean position/orientation (MPO) method* [162].

We start with the PBO method. Here, one defines for $1 \leq i \leq N$ the value $\omega_i := \|x_i - x_0\|_2$, and $D := \text{diag}(\omega_1^2, \dots, \omega_N^2) \in \mathbb{R}^{N \times N}$. Furthermore, one sets $Q = (M_1, \dots, M_N) \in \mathbb{R}^{3 \times N}$. The idea is now to use source positions close to the true source position and combine them in such a way that they match the true dipole orientation. This is expressed in the optimization problem

$$c_{\text{PBO}} := \arg \min_{c \in \mathbb{R}^N} \sum_{i=1}^N \omega_i^2 c_i^2 \quad \text{subject to } Q \cdot c = M. \quad (2.28)$$

By the method of Lagrange multipliers, it follows that at a minimizer of the objective function $f(c) = \sum_{i=1}^N \omega_i^2 c_i^2$ the gradient of f must be contained in the normal space of the manifold defined by $\{Q \cdot c = M\}$, which yields

$$2Dc = Q^\top \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \end{pmatrix},$$

which together with $Q \cdot c = M$ leads to

$$\begin{pmatrix} D & Q^\top \\ Q & 0 \end{pmatrix} \cdot \begin{pmatrix} c \\ -\frac{1}{2}\lambda_1 \\ -\frac{1}{2}\lambda_2 \\ -\frac{1}{2}\lambda_3 \end{pmatrix} = \begin{pmatrix} 0 \\ M \end{pmatrix}. \quad (2.29)$$

The PBO approach now consists of computing the vector c as the solution of the equation (2.29).

In [15], [162], and [133], it is now claimed that the left-hand side matrix of equation (2.29) is invertible. We will now show that this is, in fact, not necessarily the case.

Lemma 2.5.8. *We use the notation defined in the preceding paragraphs. Then there are combinations of dipole and mesh node positions such that the matrix*

$$\begin{pmatrix} D & Q^\top \\ Q & 0 \end{pmatrix}$$

is not invertible.

Proof. Let T denote a tetrahedron with corners p_1, p_2, p_3 , and p_4 . Let $x_0 \in T$ have the affine representation $x = \sum_{i=1}^4 \lambda_i p_i$ with $\sum_{i=1}^4 \lambda_i = 1$. For $1 \leq i \leq 4$, let q_i denote the vertex of the tetrahedron sharing the face opposite of p_i with T that is not contained in T . If $0 < \lambda_1, \lambda_2, \lambda_3, \lambda_4 < \frac{1}{2}$, e.g. when $\lambda_1 = \lambda_2 = \lambda_3 = \lambda_4 = \frac{1}{4}$, it is possible that for $1 \leq i \leq 4$ we have

$$q_i = 2 \cdot x_0 - p_i.$$

But then, the face function associated to the pair (p_i, q_i) is assigned a dipole with position $\frac{p_i + q_i}{2} = x_0$, and hence $\|\frac{p_i + q_i}{2} - x_0\| = 0$. This implies that D has four zeros on the diagonal, and hence $\text{rank}(D) = N - 4$. This implies that $\text{rank}((D, Q^\top)) \leq N - 1$, which in turn implies that

$$\text{rank} \begin{pmatrix} D & Q^\top \\ Q & 0 \end{pmatrix} \leq N + 2,$$

which since $\begin{pmatrix} D & Q^\top \\ Q & 0 \end{pmatrix} \in \mathbb{R}^{(N+3) \times (N+3)}$ implies that $\begin{pmatrix} D & Q^\top \\ Q & 0 \end{pmatrix}$ is not invertible. \square

We want to note that even though we cannot guarantee the invertibility of $\begin{pmatrix} D & Q^\top \\ Q & 0 \end{pmatrix}$ in general, we still typically expect the matrix to be invertible. For example, if for each $1 \leq i \leq N$ we have that $\omega_i^2 \neq 0$, it follows that D is invertible, and we can compute

$$\begin{pmatrix} E_N & 0 \\ -QD^{-1} & E_3 \end{pmatrix} \cdot \begin{pmatrix} D & Q^\top \\ Q & 0 \end{pmatrix} = \begin{pmatrix} D & Q^\top \\ 0 & -QD^{-1}Q^\top \end{pmatrix}.$$

It is straightforward to see that $\text{rank}(Q) = 3$, and hence since D^{-1} is positive definite, it follows that $\text{rank}(-QD^{-1}Q^\top) = 3$. Thus, the right-hand side of the preceding equation is invertible, and then also $\begin{pmatrix} D & Q^\top \\ Q & 0 \end{pmatrix}$. We thus see that $\omega_i^2 \neq 0$ for $1 \leq i \leq N$ is sufficient for the PBO matrix to be invertible. We suspect that it is also possible to show the invertibility of the PBO matrix under even weaker assumptions, but we will not try to do so in this thesis.

We now continue with the MPO method. In this method, which was first introduced in [162], one tries to select the interpolation coefficients c_1, \dots, c_N in such a way that the set of 12 equations

$$\begin{aligned} \sum_{i=1}^N c_i M_i &= M \\ \frac{1}{\alpha^{\text{ref}}} \cdot \sum_{i=1}^N c_i \langle x_i - x_0, e_j \rangle M_i &= 0 \quad \text{for } j \in \{1, 2, 3\} \end{aligned}$$

are fulfilled, where α^{ref} is a weighting parameter and e_j denotes the j -th unit vector.⁸ If we again define $Q = (M_1, \dots, M_N) \in \mathbb{R}^{3 \times N}$, and additionally for $1 \leq j \leq 3$ set

$$P_j := \frac{1}{\alpha^{\text{ref}}} \cdot \text{diag}(\langle x_1 - x_0, e_j \rangle, \dots, \langle x_N - x_0, e_j \rangle) \in \mathbb{R}^{N \times N},$$

we see that these 12 equations can be written in matrix form as

$$W := \begin{pmatrix} Q \\ Q \cdot P_1 \\ Q \cdot P_2 \\ Q \cdot P_3 \end{pmatrix} \cdot c = \begin{pmatrix} M \\ 0 \\ 0 \\ 0 \end{pmatrix},$$

⁸Note that these equations arise from the condition that $\int_{\Omega} (M \cdot \delta_{x_0}) \cdot v \, dV = \int_{\Omega} \left(\sum_{i=1}^N c_i \cdot M_i \cdot \delta_{x_i} \right) \cdot v \, dV$ for $v \in \{1, (x - x_0)^{e_1}, (x - x_0)^{e_2}, (x - x_0)^{e_3}\}$. We thus see that the MPO performs the interpolation of the dipole by matching integrals against a selection of test functions, which is completely analogous to the basic idea of the Venant approach presented in equation (2.21).

with $W \in \mathbb{R}^{12 \times N}$. In [162], one then defines $c_{\text{MPO}} = W^+ \cdot \begin{pmatrix} M \\ 0 \\ 0 \\ 0 \end{pmatrix}$, where W^+ denotes the Moore-Penrose pseudoinverse.

In [162], it is claimed that the MPO approach is superior to the PBO approach as long as one uses the full set of (up to) 22 edge and face functions defined above. In later publications of the same work group, such as [199], it is, however, claimed that MPO is less stable than PBO, and that PBO might thus be the preferred approach.

To form our own opinion, we carried out numerical comparisons of the PBO and MPO approaches, whose results can be found in section 2.8.3.⁹ During these numerical investigations, we noticed that the MPO matrix $W \in \mathbb{R}^{12 \times N}$ always had rank at most 10. This contradicts a statement in [162], where the authors claim that MPO is superior to PBO if one uses the full 22 sources, but PBO becomes superior if the MPO equation system happens to be overdetermined. But if the rank of the MPO matrix is always at most 10, MPO always tries to fulfill 12 conditions with at most 10 degrees of freedom, i.e., MPO is always overdetermined. Note, however, that this might explain the reported instability of MPO, as for a rank-deficient W the equation

$W \cdot c = \begin{pmatrix} M \\ 0 \\ 0 \\ 0 \end{pmatrix}$ is not solvable in general. But if the first three rows of this linear equation are

not fulfilled, the interpolated vector field has a wrong dipole moment, which has the potential to have a strong impact on the simulation accuracy.

Mathematically proving that the MPO matrix is always rank-deficient turned out to be surprisingly difficult. We will present the corresponding argument now.

Theorem 2.5.9. *Let*

$$W = \begin{pmatrix} Q \\ Q \cdot P_1 \\ Q \cdot P_2 \\ Q \cdot P_3 \end{pmatrix} \in \mathbb{R}^{12 \times N}$$

denote the MPO interpolation matrix as defined in the preceding paragraphs. Then

$$\text{rank}(W) \leq 10.$$

Proof. Each of the up to 22 vector functions ψ_i , $1 \leq i \leq N$, corresponds to a column in the matrix W . If x_i denotes the dipole position associated to ψ_i , and M_i denotes the dipole moment

⁹For our numerical simulations of the PBO and MPO approaches, we used their implementations in the DUNEuro toolbox, which is available here https://gitlab.dune-project.org/duneuro/duneuro/-/blob/master/duneuro/eeg/whitney_source_model.hh?ref_type=heads. We want to thank Tuuli Miinalainen and Andreas Westhoff for providing this excellent implementation, which we only slightly updated to include the most recent additions from [133].

associated to ψ_i , the corresponding column w_i of W is given by

$$w_i = \begin{pmatrix} (M_i)_1 \\ (M_i)_2 \\ (M_i)_3 \\ \frac{(x_i - x_0)_1}{\alpha^{\text{ref}}} \cdot (M_i)_1 \\ \frac{(x_i - x_0)_1}{\alpha^{\text{ref}}} \cdot (M_i)_2 \\ \frac{(x_i - x_0)_1}{\alpha^{\text{ref}}} \cdot (M_i)_3 \\ \frac{(x_i - x_0)_2}{\alpha^{\text{ref}}} \cdot (M_i)_1 \\ \frac{(x_i - x_0)_2}{\alpha^{\text{ref}}} \cdot (M_i)_2 \\ \frac{(x_i - x_0)_2}{\alpha^{\text{ref}}} \cdot (M_i)_3 \\ \frac{(x_i - x_0)_3}{\alpha^{\text{ref}}} \cdot (M_i)_1 \\ \frac{(x_i - x_0)_3}{\alpha^{\text{ref}}} \cdot (M_i)_2 \\ \frac{(x_i - x_0)_3}{\alpha^{\text{ref}}} \cdot (M_i)_3 \end{pmatrix}.$$

Now, let T denote the tetrahedron containing the source position x_0 . Let p_1, p_2, p_3 , and p_4 denote the vertices of T . Then, we get 6 columns in W for the edge functions of T , and 4 columns of W for the face functions of T , for a combined set of 10 columns. We will now show that the remaining up to 12 columns associated to exterior edges can be constructed as linear combinations of the columns corresponding to the 6 interior edges and the 4 tetrahedral faces.

To this end, let q_1 denote the unique vertex of a tetrahedron sharing a face with T that is not contained in T . Pick an arbitrary edge of the tetrahedron containing q_1 that is not an edge of T . WLOG, we can assume that this edge connects the vertices q_1 and p_3 . Furthermore, we can WLOG assume that the shared face of the two tetrahedrons under consideration has the vertices p_1, p_2 , and p_3 , and that w_1 is the column corresponding to the face function given by the pair (p_4, q_1) , w_2 is the column corresponding to the edge given by (p_4, p_3) , w_3 is the column given by the edge (p_1, p_3) , w_4 is the column given by the edge (p_4, p_1) , w_5 is the column given by the edge (p_2, p_3) , and w_6 is the column given by the edge (p_4, p_2) . Now, let $w_{\text{target}} := w_{i_0}$, where i_0 is the index corresponding to the column representing the edge from p_3 to q_1 .

We now want to show that w_{target} can be represented as a linear combination of w_1, \dots, w_6 . First, for $1 \leq i \leq N$ let (p_1^i, p_2^i) denote the pair of mesh vertices determining the function ψ_i . Then, as defined above we have $x_i = \frac{p_2^i + p_1^i}{2}$ and $M_i = \frac{p_2^i - p_1^i}{\|p_2^i - p_1^i\|}$. We then for $1 \leq i \leq N$ define $\hat{w}_i := \|p_2^i - p_1^i\| \cdot w_i$, and $\hat{w}_{\text{target}} := \hat{w}_{i_0}$, where i_0 is again the index of the column representing the edge from p_3 to q_1 . Notice that w_{target} can be represented as a linear combination of w_1, \dots, w_6 if, and only if, \hat{w}_{target} can be represented as a linear combination of $\hat{w}_1, \dots, \hat{w}_6$, and it thus suffices to show the latter statement.

For q_1 we now have a unique representation in affine coordinates as $q_1 = \sum_{k=1}^4 \lambda_k p_k$ with $\sum_{k=1}^4 \lambda_k = 1$. We now define

$$\alpha := -\lambda_1 \tag{2.30}$$

$$\beta := -\lambda_2 \tag{2.31}$$

$$\gamma := -1 - \alpha - \beta. \tag{2.32}$$

We will now show that

$$\hat{w}_{\text{target}} = \hat{w}_1 + \gamma \cdot \hat{w}_2 + \alpha \cdot (\hat{w}_3 + \hat{w}_4) + \beta \cdot (\hat{w}_5 + \hat{w}_6), \tag{2.33}$$

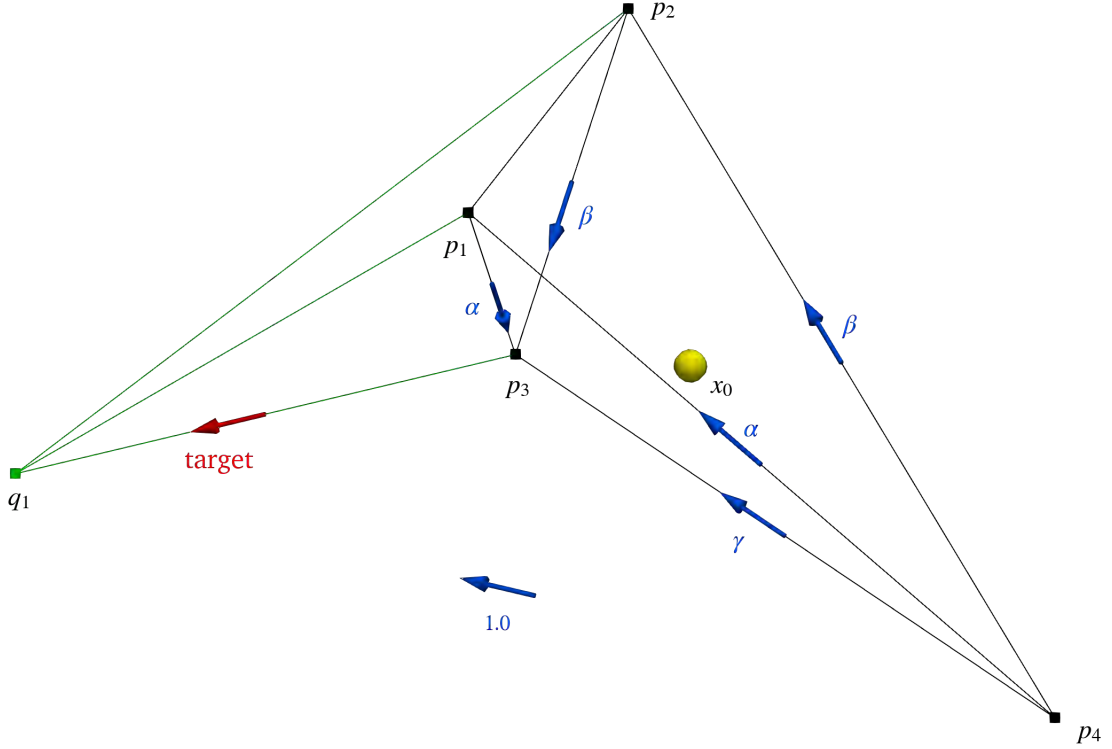


Figure 2.5: Visualization of the construction showing the linear dependence of the MPO matrix columns corresponding to exterior edges. Each relevant column is visualized by plotting its associated dipole at its position, pointing in the direction of its orientation. The position $x_0 \in T$ of the dipole one aims to interpolate is visualized by a yellow sphere.

which will finish the proof. This construction is illustrated in figure 2.5.

We first show equality in the first three rows of (2.33). To this end, note that (2.32) implies $\alpha + \beta + \gamma = -1$, and thus

$$\begin{aligned} q_1 - p_4 + \gamma \cdot (p_3 - p_4) + \alpha \cdot ((p_3 - p_1) + (p_1 - p_4)) + \beta \cdot ((p_3 - p_2) + (p_2 - p_4)) \\ = q_1 - p_4 + (\alpha + \beta + \gamma) \cdot (p_3 - p_4) = q_1 - p_3. \end{aligned}$$

Notice that this is exactly the claim of the first three rows of (2.33). The remaining nine rows are somewhat more involved. First of all, we write these equations out explicitly. Note that, after multiplication by α^{ref} , equations 4 up to 12 of (2.33) are equivalent to

$$\begin{aligned} (q_1 - p_3)_i \left(\frac{q_1 + p_3}{2} - x_0 \right)_k &= (q_1 - p_4)_i \left(\frac{q_1 + p_4}{2} - x_0 \right)_k + \gamma \cdot (p_3 - p_4)_i \left(\frac{p_3 + p_4}{2} - x_0 \right)_k \\ &\quad + \alpha \cdot (p_3 - p_1)_i \left(\frac{p_3 + p_1}{2} - x_0 \right)_k + \alpha \cdot (p_1 - p_4)_i \left(\frac{p_1 + p_4}{2} - x_0 \right)_k \\ &\quad + \beta \cdot (p_3 - p_2)_i \left(\frac{p_3 + p_2}{2} - x_0 \right)_k + \beta \cdot (p_2 - p_4)_i \left(\frac{p_2 + p_4}{2} - x_0 \right)_k. \end{aligned}$$

for $1 \leq i, k \leq 3$. As a first simplification, define $\widehat{p}_i := p_i - x_0$, $\widehat{q}_1 := q_1 - x_0$, and multiply by 2. Then, the equations above are equivalent to

$$\begin{aligned} (\widehat{q}_1 - \widehat{p}_3)_i (\widehat{q}_1 + \widehat{p}_3)_k &= (\widehat{q}_1 - \widehat{p}_4)_i (\widehat{q}_1 + \widehat{p}_4)_k + \gamma \cdot (\widehat{p}_3 - \widehat{p}_4)_i (\widehat{p}_3 + \widehat{p}_4)_k \\ &\quad + \alpha \cdot (\widehat{p}_3 - \widehat{p}_1)_i (\widehat{p}_3 + \widehat{p}_1)_k + \alpha \cdot (\widehat{p}_1 - \widehat{p}_4)_i (\widehat{p}_1 + \widehat{p}_4)_k \\ &\quad + \beta \cdot (\widehat{p}_3 - \widehat{p}_2)_i (\widehat{p}_3 + \widehat{p}_2)_k + \beta \cdot (\widehat{p}_2 - \widehat{p}_4)_i (\widehat{p}_2 + \widehat{p}_4)_k. \end{aligned}$$

for $1 \leq i, k \leq 3$. Inserting $\gamma = -1 - \alpha - \beta$, this is equivalent to

$$\begin{aligned} (\widehat{q}_1 - \widehat{p}_3)_i (\widehat{q}_1 + \widehat{p}_3)_k - (\widehat{q}_1 - \widehat{p}_4)_i (\widehat{q}_1 + \widehat{p}_4)_k + (\widehat{p}_3 - \widehat{p}_4)_i (\widehat{p}_3 + \widehat{p}_4)_k \\ = \alpha \cdot ((\widehat{p}_3 - \widehat{p}_1)_i (\widehat{p}_3 + \widehat{p}_1)_k + (\widehat{p}_1 - \widehat{p}_4)_i (\widehat{p}_1 + \widehat{p}_4)_k - (\widehat{p}_3 - \widehat{p}_4)_i (\widehat{p}_3 + \widehat{p}_4)_k) \\ + \beta \cdot ((\widehat{p}_3 - \widehat{p}_2)_i (\widehat{p}_3 + \widehat{p}_2)_k + (\widehat{p}_2 - \widehat{p}_4)_i (\widehat{p}_2 + \widehat{p}_4)_k - (\widehat{p}_3 - \widehat{p}_4)_i (\widehat{p}_3 + \widehat{p}_4)_k) \end{aligned}$$

for $1 \leq i, k \leq 3$. Multiplying out the braces leads to

$$\begin{aligned} (\widehat{p}_3 - \widehat{p}_1)_i (\widehat{p}_3 + \widehat{p}_1)_k + (\widehat{p}_1 - \widehat{p}_4)_i (\widehat{p}_1 + \widehat{p}_4)_k - (\widehat{p}_3 - \widehat{p}_4)_i (\widehat{p}_3 + \widehat{p}_4)_k \\ = (\widehat{p}_1 - \widehat{p}_4)_k \cdot (\widehat{p}_3 - \widehat{p}_4)_i - (\widehat{p}_1 - \widehat{p}_4)_i \cdot (\widehat{p}_3 - \widehat{p}_4)_k \end{aligned}$$

and

$$\begin{aligned} (\widehat{p}_3 - \widehat{p}_2)_i (\widehat{p}_3 + \widehat{p}_2)_k + (\widehat{p}_2 - \widehat{p}_4)_i (\widehat{p}_2 + \widehat{p}_4)_k - (\widehat{p}_3 - \widehat{p}_4)_i (\widehat{p}_3 + \widehat{p}_4)_k \\ = (\widehat{p}_2 - \widehat{p}_4)_k \cdot (\widehat{p}_3 - \widehat{p}_4)_i - (\widehat{p}_2 - \widehat{p}_4)_i \cdot (\widehat{p}_3 - \widehat{p}_4)_k \end{aligned}$$

and

$$\begin{aligned} (\widehat{q}_1 - \widehat{p}_3)_i (\widehat{q}_1 + \widehat{p}_3)_k + (\widehat{p}_3 - \widehat{p}_4)_i (\widehat{p}_3 + \widehat{p}_4)_k - (\widehat{q}_1 - \widehat{p}_4)_i (\widehat{q}_1 + \widehat{p}_4)_k \\ = (\widehat{q}_1 - \widehat{p}_4)_k (\widehat{p}_4 - \widehat{p}_3)_i - (\widehat{q}_1 - \widehat{p}_4)_i (\widehat{p}_4 - \widehat{p}_3)_k. \end{aligned}$$

We thus see that the set of equations above is equivalent to

$$\begin{aligned} (\widehat{q}_1 - \widehat{p}_4)_k (\widehat{p}_4 - \widehat{p}_3)_i - (\widehat{q}_1 - \widehat{p}_4)_i (\widehat{p}_4 - \widehat{p}_3)_k \\ = \alpha \cdot ((\widehat{p}_1 - \widehat{p}_4)_k \cdot (\widehat{p}_3 - \widehat{p}_4)_i - (\widehat{p}_1 - \widehat{p}_4)_i \cdot (\widehat{p}_3 - \widehat{p}_4)_k) \\ + \beta \cdot ((\widehat{p}_2 - \widehat{p}_4)_k \cdot (\widehat{p}_3 - \widehat{p}_4)_i - (\widehat{p}_2 - \widehat{p}_4)_i \cdot (\widehat{p}_3 - \widehat{p}_4)_k) \end{aligned}$$

for $1 \leq i, k \leq 3$. Now, by construction we have $q_1 = \sum_{j=1}^4 \lambda_j p_j$ with $\sum_{j=1}^4 \lambda_j = 1$. But then, we also have

$$\widehat{q}_1 = q_1 - x_0 = \sum_{j=1}^4 \lambda_j (p_j - x_0) = \sum_{j=1}^4 \lambda_j \widehat{p}_j,$$

which in turn yields

$$\widehat{q}_1 - \widehat{p}_4 = \sum_{j=1}^3 \lambda_j (\widehat{p}_j - \widehat{p}_4).$$

But now, we have finally reduced the set of equations to a form where we can show that they

are true. Starting from the left-hand side, we have for $1 \leq i, k \leq 3$ that

$$\begin{aligned}
& (\hat{q}_1 - \hat{p}_4)_k (\hat{p}_4 - \hat{p}_3)_i - (\hat{q}_1 - \hat{p}_4)_i (\hat{p}_4 - \hat{p}_3)_k \\
&= \sum_{j=1}^3 \lambda_j (\hat{p}_j - \hat{p}_4)_k (\hat{p}_4 - \hat{p}_3)_i - \sum_{j=1}^3 \lambda_j (\hat{p}_j - \hat{p}_4)_i (\hat{p}_4 - \hat{p}_3)_k \\
&= \lambda_1 \cdot ((\hat{p}_1 - \hat{p}_4)_k (\hat{p}_4 - \hat{p}_3)_i - (\hat{p}_1 - \hat{p}_4)_i (\hat{p}_4 - \hat{p}_3)_k) + \lambda_2 \cdot ((\hat{p}_2 - \hat{p}_4)_k (\hat{p}_4 - \hat{p}_3)_i - (\hat{p}_2 - \hat{p}_4)_i (\hat{p}_4 - \hat{p}_3)_k) \\
&\quad + \lambda_3 \cdot \underbrace{((\hat{p}_3 - \hat{p}_4)_k (\hat{p}_4 - \hat{p}_3)_i - (\hat{p}_3 - \hat{p}_4)_i (\hat{p}_4 - \hat{p}_3)_k)}_{=0} \\
&\stackrel{(2.30) \text{ and } (2.31)}{=} \alpha \cdot ((\hat{p}_1 - \hat{p}_4)_k (\hat{p}_3 - \hat{p}_4)_i - (\hat{p}_1 - \hat{p}_4)_i (\hat{p}_3 - \hat{p}_4)_k) \\
&\quad + \beta \cdot ((\hat{p}_2 - \hat{p}_4)_k (\hat{p}_3 - \hat{p}_4)_i - (\hat{p}_2 - \hat{p}_4)_i (\hat{p}_3 - \hat{p}_4)_k).
\end{aligned}$$

Notice that this is exactly the set of equations for which we have shown above that they are equivalent to the set of equations arising from rows 4 to 12 of (2.33). We have thus shown that equation (2.33) is valid, which finishes the proof. \square

This finishes our discussion of the established direct approaches for the EEG forward problem, namely the *partial integration approach* (see definition 2.5.1), the *Venant approach* (see definition 2.5.3), and the *Whitney approach* (see definition 2.5.6). This, finally, brings us to subtraction-type approaches.

2.5.2 Subtraction approaches for the EEG forward problem

One of the central difficulties in the EEG forward problem is how to handle the dipole singularity. In the direct approaches, this is addressed by either ignoring it (in the partial integration approach, leading to relatively large errors), or by substituting the singular dipole, resp. its divergence, with a more regular object (e.g., a distributed monopole cloud in the Venant approach, or a combination of $H(\text{div}; \Omega)$ functions in the Whitney approach). Subtraction approaches deal with this problem in a different way. The idea is to “subtract” the singularity out of the problem formulation, which leads to a more regular problem. This more regular problem can then be solved in the standard way, and the singularity can be added back in in a post-processing step. This idea is in fact quite old (for the standards of bioelectromagnetism research), and has been deeply investigated both theoretically and numerically over a large number of publications [201, 197, 9, 126, 182, 239, 48, 53, 20, 18]. The general principle is as follows.

While we cannot analytically compute the electric potential u of a dipole in a general head model Ω , we can compute this potential in certain simplified geometries. In particular, we can analytically compute the electric potential for an unbounded volume conductor with homogeneous conductivity. Concretely, if $\sigma^\infty \in \mathbb{R}^{3 \times 3}$ is positive definite, then we have a simple analytical formula for a function $u_{x_0, M, \sigma^\infty}^\infty$ such that

$$\text{div}(\sigma^\infty \nabla u_{x_0, M, \sigma^\infty}^\infty) = \text{div}(M \cdot \delta_{x_0}) \quad \text{on } \mathbb{R}^3. \quad (2.34)$$

in the distributional sense for $x_0, M \in \mathbb{R}^3$. In section 2.6.2 of this thesis, we will show how to compute this $u_{x_0, M, \sigma^\infty}^\infty$, with the final result summarized in theorem 2.6.6. For the remainder of this section, we will always assume that we want to simulate the dipolar potential due to a fixed dipole at position $x_0 \in \Omega$ with moment $M \in \mathbb{R}^3$ and $\sigma^\infty := \sigma(x_0)$, and will hence simply write u^∞ for $u_{x_0, M, \sigma^\infty}^\infty$. The idea of the aforementioned papers is now to not directly solve for

the head potential u , but instead solve for the deviation of u from the potential in an unbounded homogeneous medium, i.e., for $u_{\text{classical}}^c := u - u^\infty$.

For this first introduction, we will focus on the PEM model. Then, according to definition 2.1.1, the electric potential u is given by

$$\begin{aligned} \operatorname{div}(\sigma \nabla u) &= \operatorname{div}(j^P) && \text{in } \Omega \\ \langle \sigma \nabla u, \eta \rangle &= 0 && \text{on } \partial\Omega \end{aligned}$$

Now, let $\sigma^c := \sigma - \sigma^\infty$. We assume that σ is constant on a small environment of x_0 , which implies that $\sigma^c = 0$ on this environment. For $u_{\text{classical}}^c$, we then get

$$\begin{aligned} \operatorname{div}(\sigma \nabla u_{\text{classical}}^c) &= \operatorname{div}(\sigma \nabla u) - \operatorname{div}(\sigma \nabla u^\infty) \\ &= \operatorname{div}(j^P) - \underbrace{\operatorname{div}(\sigma^\infty \nabla u^\infty)}_{=\operatorname{div}(j^P)} - \operatorname{div}(\sigma^c \nabla u^\infty) \\ &= -\operatorname{div}(\sigma^c \nabla u^\infty), \end{aligned}$$

and we see that $u_{\text{classical}}^c$ is described by the problem

$$\begin{aligned} \operatorname{div}(\sigma \nabla u_{\text{classical}}^c) &= -\operatorname{div}(\sigma^c \nabla u^\infty) && \text{in } \Omega \\ \langle \sigma \nabla u_{\text{classical}}^c, \eta \rangle &= -\langle \sigma \nabla u^\infty, \eta \rangle && \text{on } \partial\Omega. \end{aligned}$$

Proceeding in the standard way by multiplying with a test function φ and applying partial integration yields

$$\begin{aligned} \int_{\Omega} \langle \sigma \nabla u_{\text{classical}}^c, \nabla \varphi \rangle dV &= \int_{\partial\Omega} \langle \sigma \nabla u_{\text{classical}}^c, \eta \rangle \cdot \varphi dS - \int_{\Omega} \operatorname{div}(\sigma \nabla u_{\text{classical}}^c) \cdot \varphi dV \\ &= - \int_{\partial\Omega} \langle \sigma \nabla u^\infty, \eta \rangle \cdot \varphi dS + \int_{\Omega} \operatorname{div}(\sigma^c \nabla u^\infty) \varphi dV \\ &= - \int_{\partial\Omega} \langle \sigma \nabla u^\infty, \eta \rangle \cdot \varphi dS + \int_{\partial\Omega} \langle \sigma^c \nabla u^\infty, \eta \rangle \cdot \varphi dS - \int_{\Omega} \langle \sigma^c \nabla u^\infty, \nabla \varphi \rangle dV \\ &= - \int_{\partial\Omega} \langle \sigma^\infty \nabla u^\infty, \eta \rangle \cdot \varphi dS - \int_{\Omega} \langle \sigma^c \nabla u^\infty, \nabla \varphi \rangle dV. \end{aligned}$$

This finishes the derivation of the PEM subtraction approach.

Definition 2.5.10 (EEG Subtraction approach, PEM case). Assume we have a current dipole at position $x_0 \in \Omega$ with moment $M \in \mathbb{R}^3$. Let $V \subset H^1(\Omega)$ be a closed subspace. Let a_{PEM} be defined as in (2.12). Then, the PEM subtraction approach is given by the problem of finding $u \in V$ such that

$$a_{\text{PEM}}(u, v) = - \int_{\partial\Omega} \langle \sigma^\infty \nabla u^\infty, \eta \rangle \cdot \varphi dS - \int_{\Omega} \langle \sigma^c \nabla u^\infty, \nabla \varphi \rangle dV$$

for all $v \in V$.

Remark 2.5.11. We want to emphasize that, in contrast to the definitions in section 2.5.1, in definition 2.5.10 the choice $V = H^1(\Omega)$ is explicitly allowed. In fact, the subtraction approach can be used to show that the continuous EEG forward problem has a solution in the first place. If we denote

$$l_S : H^1(\Omega) \rightarrow \mathbb{R}; v \mapsto - \int_{\partial\Omega} \langle \sigma^\infty \nabla u^\infty, \eta \rangle \cdot \varphi dS - \int_{\Omega} \langle \sigma^c \nabla u^\infty, \nabla \varphi \rangle dV, \quad (2.35)$$

the well-definedness of l_S follows from $\sigma^c = 0$ on an environment of x_0 and the fact that u^∞ is smooth and bounded on $\mathbb{R}^3 \setminus B_\epsilon(x_0)$, where $B_\epsilon(x_0)$ denotes the ball of radius $\epsilon > 0$ around x_0 . It then follows from the trace theorem that l_S defines a continuous linear functional on $H^1(\Omega)$. If we can now show that $l_S(1) = 0$, it follows that for all closed $V \subset H^1(\Omega)$ the subtraction approach has a solution that is unique up to an additive constant, using the same proof as in theorem 2.4.12. The statement $l_S(1) = 0$ is in fact true, as is e.g. shown in [239], Theorem 3.7. Later on, we will generalize the subtraction approach to a larger class of approaches, called *local subtraction* approaches, and will then, in the proof of theorem 2.5.13, show that each of the corresponding right-hand side functionals vanishes on constant functions. The statement $l_S(1) = 0$ will then follow as a special case of that theorem, and we will not give a separate proof here.

We thus see that the subtraction approach has a mathematically rigorous foundation. Furthermore, it has been shown that the subtraction approach is able to produce highly accurate EEG simulations, see e.g. [116]. Nevertheless, if one looks at the standard open source toolboxes in the bioelectromagnetism community, i.e., Brainstorm [129] and FieldTrip [219], both, at least at the time of writing this thesis, choose the Venant approach from definition 2.5.3 as the default strategy for EEG forward modeling.¹⁰ The reason for this is the computational demand of the subtraction approach. To assemble the FEM right-hand side, one needs to evaluate the values $l_S(\varphi_1), \dots, l_S(\varphi_N)$ for a basis $\{\varphi_1, \dots, \varphi_N\}$ of the test function space. Looking at (2.35), one sees that many of these values will be non-zero, and assembling the subtraction right-hand side effectively amounts to iterating over the complete head domain and assembling integrals on each single element. For the EEG and MEG inverse problems, one typically needs to solve the forward problems for tens of thousands of sources, and performing an integration over the whole head domain for each of these sources quickly becomes computationally infeasible. The direct approaches, on the other hand, be it partial integration, Venant, or Whitney, all produce sparse right-hand sides with at most a couple of dozen non-zero entries in the right-hand side. Using e.g. the DUNEuro-toolbox [186], these can then be assembled rapidly, i.e., in a matter of milliseconds on contemporary hardware, while the corresponding assembly of the subtraction right-hand sides can take hours, or even days. Taking the Venant approach as a default was thus a practical necessity.

We want to emphasize that, even though the current implementation of the subtraction approach in DUNEuro does this, it is typically not necessary to iterate over the whole head volume to assemble the subtraction right-hand side. Assume that σ^∞ is isotropic. Then, u^∞ is harmonic on $\mathbb{R}^3 \setminus \{x_0\}$. Now, let $\varphi \in H^1(\Omega)$ be a continuous test function that is only non-zero on a region $\hat{\Omega} \subset \Omega$ where σ is constant and isotropic, and assume that φ is zero on the boundary $\partial\Omega$. We claim that then $l_S(\varphi) = 0$. If $\sigma^c = 0$, this is clear. Otherwise, we can on $\hat{\Omega}$ write $\sigma^c = \lambda \cdot E_3$ with $\lambda \neq 0$ and E_3 the identity matrix, and we can compute

$$\begin{aligned} l_S(\varphi) &= - \int_{\hat{\Omega}} \langle \sigma^c \nabla u^\infty, \nabla \varphi \rangle dV = -\lambda \cdot \int_{\hat{\Omega}} \langle \nabla u^\infty, \nabla \varphi \rangle dV \\ &= -\lambda \cdot \int_{\partial\hat{\Omega}} \langle \nabla u^\infty, \eta_{\hat{\Omega}} \rangle \cdot \underbrace{\varphi}_{=0} dS + \lambda \cdot \int_{\hat{\Omega}} \underbrace{\Delta(u^\infty)}_{=0} \cdot \varphi dV \\ &= 0. \end{aligned}$$

¹⁰Concretely, Brainstorm-DUNEuro by default uses the Nüking-Venant. FieldTrip-SimBio defaults to a slightly reweighted version of the Buchner-Venant, where the implementer for some reason decided to repeat the monopole condition three times in the matrix, see https://github.com/fieldtrip/fieldtrip/blob/master/external/simbio/sb_calc_ven_loads.m.

As a particular special case, this implies that for a tetrahedral mesh with piecewise affine trial and test functions and isotropic conductivities, only those nodal basis functions corresponding to mesh vertices lying on a tissue interface give a non-zero contribution to the FEM subtraction right-hand side, and the right-hand side can thus be assembled by iterating only over the tissue interfaces. Apart from excluding anisotropic conductivities, this “optimization” however does not substantially alter the problem of computational efficiency, as, due to the complex folding of the human brain and the nesting of different tissues in the head (see e.g. figure 2.3b), an iteration over only the tissue interfaces is not far off from an iteration over the whole head model, and will in particular not lead to an orders of magnitude reduction in computation times that would be needed to achieve an approach that is competitive in practice to the direct approaches.

On a side note, notice that for isotropic σ^∞ , the Gauss law (2.1) implies for the charge density ρ that

$$\frac{\rho}{\epsilon_0} = \operatorname{div}(E) = -\Delta u = -\Delta u_{\text{classical}}^c - \Delta u^\infty \iff -\Delta u_{\text{classical}}^c = \frac{\rho}{\epsilon_0} + \frac{\operatorname{div}(M \cdot \delta_{x_0})}{\sigma^\infty}.$$

On the other hand, under the assumption that σ is isotropic, the equation

$$\operatorname{div}(\sigma \nabla u_{\text{classical}}^c) = -\operatorname{div}(\sigma^c \nabla u^\infty)$$

implies that $-\Delta u_{\text{classical}}^c$ is zero outside of tissue interfaces, and it follows that $\frac{\rho}{\epsilon_0} + \frac{\operatorname{div}(M \cdot \delta_{x_0})}{\sigma^\infty}$ is concentrated on the tissue interfaces. From this perspective, one can say that, in the case of isotropic σ , the correction potential of the subtraction approach is the part of the potential that is due to the accumulation of charges at the tissue boundaries. This is very similar to the interpretation of the secondary electric field in the modeling of transcranial magnetic stimulation.

We thus see that the global nature of the subtraction functional l_S leads to a high computational demand, which renders the subtraction approach essentially unusable in practice. Accordingly, a lot of work has gone into trying to increase the efficiency of subtraction approaches. One such idea is the so-called *projected subtraction approach* from [239]. Here, the idea is to project the homogeneous potential u^∞ into a simpler function space. If one, e.g., uses a tetrahedral mesh with piecewise affine trial functions, each basis function φ_i is associated with a node x_i , and we can approximate u^∞ by its nodal interpolation via

$$u^\infty \approx \sum_{j=1}^N u^\infty(x_j) \cdot \varphi_j,$$

where $N \in \mathbb{N}$ denotes the number of mesh nodes. If we then define $K = (k_{ij})_{1 \leq i, j \leq N} \in \mathbb{R}^{N \times N}$ and $S = (s_{ij})_{1 \leq i, j \leq N} \in \mathbb{R}^{N \times N}$ by

$$\begin{aligned} k_{ij} &:= \int_{\Omega} \langle \sigma^c \nabla \varphi_j, \nabla \varphi_i \rangle dV \\ s_{ij} &:= \int_{\partial\Omega} \langle \sigma^\infty \nabla \varphi_j, \eta \rangle \varphi_i dS, \end{aligned}$$

we get

$$\begin{aligned} l_S(\varphi_i) &= - \int_{\partial\Omega} \langle \sigma^\infty \nabla u^\infty, \eta \rangle \cdot \varphi_i dS - \int_{\Omega} \langle \sigma^c \nabla u^\infty, \nabla \varphi_i \rangle dV \\ &\approx \langle -(S + K) \cdot \begin{pmatrix} u^\infty(x_1) \\ \vdots \\ u^\infty(x_N) \end{pmatrix}, e_i \rangle, \end{aligned}$$

where e_i denotes the i -th unit vector. Note that the matrices K and S need to be assembled only once. For the projected subtraction approach, we thus do not need to compute complicated integrals for each source of interest, but instead only evaluate the potential u^∞ and perform a matrix-vector multiplication.

The projected subtraction approach thus substitutes the true right-hand side with an approximation based on a linear approximation of the homogeneous potential. In [20], it is shown that the accuracy of this approach is highly dependent on the quality of the mesh elements, and, in particular, that the presence of strongly deformed tetrahedrons can massively degrade the accuracy of the approach. In [20], it is instead suggested to project the gradient of the homogeneous potential, and not the potential itself. Concretely, if one again assumes a tetrahedral mesh with a nodal basis $\varphi_1, \dots, \varphi_N$, associated to the mesh vertices x_1, \dots, x_N , one approximates ∇u^∞ via

$$\nabla u^\infty \approx \sum_{j=1}^N \varphi_j \cdot \nabla u^\infty(x_j).$$

If one now defines $F = (f_{ij})_{1 \leq i, j \leq N} \in \mathbb{R}^{N \times 3N}$ and $G = (g_{ij})_{1 \leq i, j \leq N} \in \mathbb{R}^{N \times 3N}$ via

$$\begin{aligned} f_{ij} &:= \int_{\Omega} \varphi_j \cdot \nabla \varphi_i^\top \cdot \sigma^c dV && \in \mathbb{R}^{1 \times 3} \\ g_{ij} &:= \int_{\partial\Omega} \varphi_j \cdot \varphi_i \cdot \eta^\top \cdot \sigma^\infty dS && \in \mathbb{R}^{1 \times 3} \end{aligned}$$

this approximation leads to

$$\begin{aligned} l_S(\varphi_i) &= - \int_{\partial\Omega} \langle \sigma^\infty \nabla u^\infty, \eta \rangle \cdot \varphi_i dS - \int_{\Omega} \langle \sigma^c \nabla u^\infty, \nabla \varphi_i \rangle dV \\ &\approx \langle -(G + F) \cdot \begin{pmatrix} \nabla u^\infty(x_1) \\ \vdots \\ \nabla u^\infty(x_N) \end{pmatrix}, e_i \rangle. \end{aligned}$$

In [20], this is called the *projected gradient subtraction* approach. Furthermore, it is shown that, for this approach, one can derive better theoretical error bounds than for the projected subtraction approach, and that the projected gradient subtraction approach also numerically outperforms the projected subtraction approach over all investigated models.

But even then, the projected gradient subtraction method remains unsatisfactory in two ways. First of all, the true source functional l_S is still substituted with an approximation, which noticeably degrades the accuracy of the approach, as is shown in [20]. Secondly, assembling the vector $(\nabla u^\infty(x_1), \dots, \nabla u^\infty(x_N))^\top$ still requires an iteration over every mesh vertex for each source under consideration, which, while being much more efficient than performing numerical integration on each mesh element, is still much slower than assembling e.g. the multipolar Venant right-hand side.

Another recent, and particularly noteworthy, contribution is given in [18]. In that paper, the author manages to derive analytical expressions for the subtraction right-hand side

$$l_S(\varphi_i) = - \int_{\partial\Omega} \langle \sigma^\infty \nabla u^\infty, \eta \rangle \cdot \varphi_i dS - \int_{\Omega} \langle \sigma^c \nabla u^\infty, \nabla \varphi_i \rangle dV$$

in the case of a tetrahedral mesh with piecewise affine trial and test functions and isotropic source conductivity σ^∞ . We will, in fact, heavily build on this work in section 2.6.3. This development massively improves subtraction approaches. First, it greatly increases the numerical stability of the approach. If we imagine a source close to a conductivity jump, we essentially have to integrate $\frac{1}{\|x\|^2}$ close to zero. Using numerical quadrature, this quickly runs into problems, but the analytical integration stays stable even for sources with a very small distance to a conductivity jump. In [18], it is, for example, demonstrated that, for meshes with standard resolutions, the analytical integration enables an accurate simulation of dipolar sources with a distance of 0.125 mm to a conductivity jump. Secondly, when compared with numerical integration, the analytical integration is quite efficient. In [18] it is shown that analytically computing the subtraction right-hand side is computationally as expensive as using numerical quadrature with an order of 2. We, however, want to emphasize that, even with the analytical integration, assembling the FEM right-hand side still requires an iteration over the whole head model, which is, in practice, not competitive with the direct approaches. Following [18], we will call a FEM simulation with an analytically computed right-hand side given by l_S the *analytical subtraction approach*. Note that the projected subtraction approach, the projected gradient subtraction approach, and the analytical subtraction approach all still try to assemble the functional given by definition 2.5.10, and it is, in my opinion, somewhat misleading to act like these are different approaches. But the literature does this, and, for the sake of clarity, we shall do the same in this thesis. For reference, computing the right-hand side given by l_S using numerical quadrature is typically called the *full subtraction approach* in the literature [48].

One central topic of this PhD thesis was investigating whether it is possible to improve the computational efficiency of subtraction approaches. This investigation resulted in the (further) development of a new approach, which we termed the *local subtraction approach*, and which was published in [91]. We now want to derive this approach.

The first steps toward the local subtraction approach were developed in the PhD thesis by Andreas Nüßing [143]. In there, the author investigates discontinuous Galerkin approaches for the EEG forward problem, and, as part of this investigation, introduces an idea for reducing the computational demand of subtraction approaches.¹¹ In this work, we develop this idea further and extend it to continuous Galerkin approaches. The central idea goes as follows. Our original goal was to eliminate the dipole singularity from the problem formulation. This is done by subtracting u^∞ from u and solving for this difference. While this cancels the singularity, it replaces the focal point dipole $\text{div}(M \cdot \delta_{x_0})$ in the equation with $-\text{div}(\sigma^c \nabla u^\infty)$, which gives non-zero right-hand side contributions at all tissue interfaces. These global non-zero contributions require an iteration over the whole volume conductor to resolve, which leads to the high computational demand, prohibiting the practical use of subtraction approaches. The central insight now is that the only purpose of subtracting u^∞ is cancelling the dipole singularity. But to achieve this, we only need to subtract u^∞ in the immediate environment of the source position, and not on the whole head domain. If we now choose some cutoff function χ in such a way that χ is 1 on an

¹¹For reference, we want to note that what we now call the “local subtraction approach” was originally termed the “localized subtraction approach” in [143]. We decided to change the name, as colleagues from the MEEG community would repeatedly confuse the approach as an inverse method, because the “localized” in the name reminded them of “source localization”.

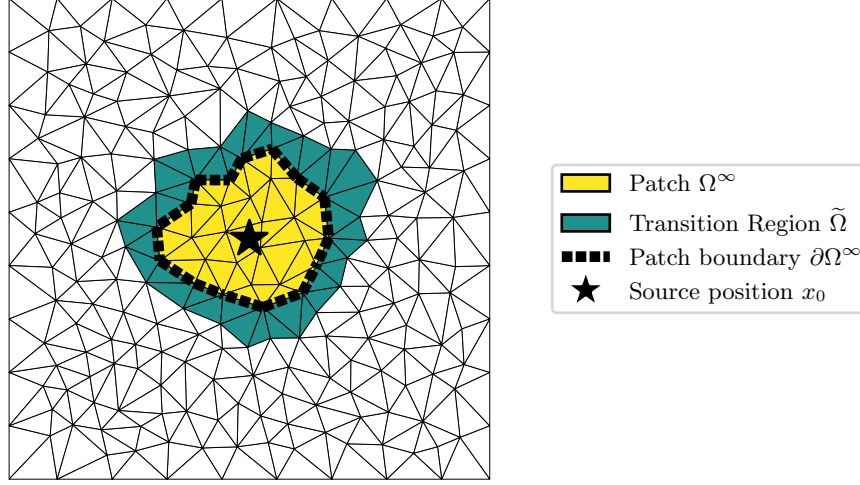


Figure 2.6: 2-dimensional visualization of the construction of the cutoff function χ , the patch Ω^∞ , and the transition region $\tilde{\Omega}$. In the yellow region, we have $\chi = 1$, in the white region we have $\chi = 0$, and χ drops from 1 to 0 in the green region. This is a modified version of Fig. 2.1 from [91].

environment of the source position, but quickly falls off to 0, we can also try to solve for

$$u^c = u - \chi \cdot u^\infty, \quad (2.36)$$

where again u denotes the head potential, and we call u^c the *correction potential*. Then, because χ is 1 in an environment of the source position, subtracting $\chi \cdot u^\infty$ still cancels the dipole singularity, but because χ is (ideally) only non-zero close to x_0 , we have that $\text{div}(\sigma \nabla u^c)$ vanishes on a large portion of the head domain, leading to a sparse right-hand side which can then be rapidly assembled.

In the discontinuous Galerkin context, the author of [143] suggested simply choosing χ as the indicator function of some homogeneous environment of the source position x_0 . Thus, in this context χ is discontinuous. If we aim for a continuous Galerkin formulation, we cannot choose χ this way. For readers interested in the discontinuous case, we refer to [143]. For the remainder of this thesis, we will focus on the continuous Galerkin case.

Concretely, we will require that $\chi \in H^1(\Omega)$. Additionally, let $\Omega^\infty \subset \Omega$ denote an region with $x_0 \in \Omega^\infty$ and such that $\chi = 1$ on Ω^∞ . Furthermore, let $\tilde{\Omega} \subset \Omega$ be such that $\chi = 0$ on $\Omega \setminus (\Omega^\infty \cup \tilde{\Omega})$ and $\Omega^\infty \cap \tilde{\Omega} = \emptyset$. We call Ω^∞ the *patch* and $\tilde{\Omega}$ the *transition region*, as χ drops of from 1 to 0 here. This construction is illustrated in figure 2.6.

Note that Ω^∞ has a different meaning here than e.g. in [239]. In that paper, Ω^∞ is used to denote a region where $\sigma^c = 0$. In this thesis, Ω^∞ is used to denote a region where the cutoff function χ is equal to 1, and can, in particular, contain multiple tissue compartments. In our notation, we thus in general have $\sigma^c \neq 0$ on Ω^∞ . Furthermore, even though we, just as [239], assume that σ is constant on a small environment of the source position x_0 , we do not associate a name to this environment, as it only implicitly enters the final problem description in establishing that the corresponding functional is well-defined, as was the case in definition 2.5.10.

Our goal is now to derive a weak formulation for $u^c = u - \chi \cdot u^\infty$. We aim to derive weak formulations both for the PEM model from definition 2.1.1 as well as for the CEM model from definition 2.1.2. Our derivation closely follows the one given in [91]. We start with

$$\begin{aligned} \operatorname{div}(\sigma \nabla u^c) &= \operatorname{div}(\sigma \nabla u) - \operatorname{div}(\sigma \nabla \chi u^\infty) = \operatorname{div}(M \delta_{x_0}) - \operatorname{div}(\sigma^\infty \nabla \chi u^\infty) - \operatorname{div}(\sigma^c \nabla \chi u^\infty) \\ &= \operatorname{div}(M \delta_{x_0}) - \operatorname{div}(\sigma^\infty \nabla u^\infty) - \operatorname{div}(\sigma^\infty \nabla (\chi - 1) u^\infty) - \operatorname{div}(\sigma^c \nabla \chi u^\infty) \\ &= -\operatorname{div}(\sigma^\infty \nabla (\chi - 1) u^\infty) - \operatorname{div}(\sigma^c \nabla \chi u^\infty), \end{aligned}$$

and we thus see that u^c is given by the boundary value problem

$$\begin{aligned} \operatorname{div}(\sigma \nabla u^c) &= -\operatorname{div}(\sigma^\infty \nabla (\chi - 1) u^\infty) - \operatorname{div}(\sigma^c \nabla \chi u^\infty) && \text{in } \Omega \\ \langle \sigma \nabla u^c, \eta \rangle &= -\langle \sigma \nabla \chi \cdot u^\infty, \eta \rangle && \text{on } \partial\Omega. \end{aligned}$$

If now φ is a test function, multiplication and integration by parts yields

$$\int_{\Omega} \operatorname{div}(\sigma \nabla u^c) \varphi \, dV = \int_{\partial\Omega} \langle \sigma \nabla u^c, \eta \rangle \varphi \, dS - \int_{\Omega} \langle \sigma \nabla u^c, \nabla \varphi \rangle \, dV. \quad (2.37)$$

But, by the preceding computation, we also have

$$\begin{aligned} \int_{\Omega} \operatorname{div}(\sigma \nabla u^c) \varphi \, dV &= \int_{\Omega} (-\operatorname{div}(\sigma^\infty \nabla (\chi - 1) u^\infty) - \operatorname{div}(\sigma^c \nabla \chi u^\infty)) \varphi \, dV \\ &= -\int_{\partial\Omega} \langle \sigma^\infty \nabla (\chi - 1) u^\infty, \eta \rangle \varphi \, dS + \int_{\Omega} \langle \sigma^\infty \nabla (\chi - 1) u^\infty, \nabla \varphi \rangle \, dV \\ &\quad - \int_{\partial\Omega} \langle \sigma^c \nabla \chi u^\infty, \eta \rangle \varphi \, dS + \int_{\Omega} \langle \sigma^c \nabla \chi u^\infty, \nabla \varphi \rangle \, dV. \end{aligned} \quad (2.38)$$

Now, we have

$$\begin{aligned} -\int_{\partial\Omega} \langle \sigma^\infty \nabla (\chi - 1) u^\infty, \eta \rangle \varphi \, dS - \int_{\partial\Omega} \langle \sigma^c \nabla \chi u^\infty, \eta \rangle \varphi \, dS \\ = -\int_{\partial\Omega} \langle \sigma \nabla \chi u^\infty, \eta \rangle \varphi \, dS + \int_{\partial\Omega} \langle \sigma^\infty \nabla u^\infty, \eta \rangle \varphi \, dS. \end{aligned}$$

Using this, combining equations (2.37) and (2.38) yields

$$\begin{aligned} \int_{\Omega} \langle \sigma \nabla u^c, \nabla \varphi \rangle \, dV - \int_{\partial\Omega} \langle \sigma \nabla u, \eta \rangle \varphi \, dS \\ = -\int_{\partial\Omega} \langle \sigma^\infty \nabla u^\infty, \eta \rangle \varphi \, dS - \int_{\Omega} \langle \sigma^\infty \nabla (\chi - 1) u^\infty, \nabla \varphi \rangle \, dV - \int_{\Omega} \langle \sigma^c \nabla \chi u^\infty, \nabla \varphi \rangle \, dV. \end{aligned} \quad (2.39)$$

We will now first further investigate the right-hand side of this equation. Once we are done with this, we will specialize to either PEM or CEM by treating the second summand on the left-hand side. Now, since on Ω^∞ we have $\chi = 1$, we have

$$\int_{\Omega} \langle \sigma^\infty \nabla (\chi - 1) u^\infty, \nabla \varphi \rangle \, dV = \int_{\Omega \setminus \Omega^\infty} \langle \sigma^\infty \nabla (\chi - 1) u^\infty, \nabla \varphi \rangle \, dV$$

and

$$\int_{\Omega} \langle \sigma^c \nabla \chi u^\infty, \nabla \varphi \rangle \, dV = \int_{\Omega \setminus \Omega^\infty} \langle \sigma^c \nabla \chi u^\infty, \nabla \varphi \rangle \, dV + \int_{\Omega^\infty} \langle \sigma^c \nabla u^\infty, \nabla \varphi \rangle \, dV,$$

and thus

$$\begin{aligned} & \int_{\Omega} \langle \sigma^{\infty} \nabla(\chi - 1)u^{\infty}, \nabla \varphi \rangle dV + \int_{\Omega} \langle \sigma^c \nabla \chi u^{\infty}, \nabla \varphi \rangle dV \\ &= \int_{\Omega^{\infty}} \langle \sigma^c \nabla u^{\infty}, \nabla \varphi \rangle dV + \int_{\Omega \setminus \Omega^{\infty}} \langle \sigma \nabla \chi u^{\infty}, \nabla \varphi \rangle dV - \int_{\Omega \setminus \Omega^{\infty}} \langle \sigma^{\infty} \nabla u^{\infty}, \nabla \varphi \rangle dV \end{aligned}$$

Now, since $\chi = 0$ on $\Omega \setminus \overline{(\tilde{\Omega} \cup \Omega^{\infty})}$, we have

$$\int_{\Omega \setminus \Omega^{\infty}} \langle \sigma \nabla \chi u^{\infty}, \nabla \varphi \rangle dV = \int_{\tilde{\Omega}} \langle \sigma \nabla \chi u^{\infty}, \nabla \varphi \rangle dV.$$

Inserting this into (2.39), we arrive at

$$\begin{aligned} & \int_{\Omega} \langle \sigma \nabla u^c, \nabla \varphi \rangle dV - \int_{\partial \Omega} \langle \sigma \nabla u, \eta \rangle \varphi dS \\ &= - \int_{\Omega^{\infty}} \langle \sigma^c \nabla u^{\infty}, \nabla \varphi \rangle dV - \int_{\tilde{\Omega}} \langle \sigma \nabla \chi u^{\infty}, \nabla \varphi \rangle dV \\ & \quad + \int_{\Omega \setminus \Omega^{\infty}} \langle \sigma^{\infty} \nabla u^{\infty}, \nabla \varphi \rangle dV - \int_{\partial \Omega} \langle \sigma^{\infty} \nabla u^{\infty}, \eta \rangle \varphi dS. \quad (2.40) \end{aligned}$$

Now, let $\partial \Omega_{\text{ext}}^{\infty} := \partial \Omega^{\infty} \cap \partial \Omega$ denote the part of the boundary of the patch that coincides with the boundary of the head domain, and let $\partial \Omega_{\text{int}}^{\infty} := \partial \Omega^{\infty} \setminus \partial \Omega_{\text{ext}}^{\infty}$ denote the part of the patch boundary that lies on the inside of the head domain. Furthermore, let $\eta_{\Omega^{\infty}}$ denote the outer normal of the patch boundary, and $\eta_{\Omega \setminus \Omega^{\infty}}$ the outer normal of $\Omega \setminus \Omega^{\infty}$. Then up to a set of two-dimensional measure zero we have $\partial(\Omega \setminus \Omega^{\infty}) = (\partial \Omega \setminus \partial \Omega_{\text{ext}}^{\infty}) \cup \partial \Omega_{\text{int}}^{\infty}$, which in combination with $\text{div}(\sigma^{\infty} \nabla u^{\infty}) = 0$ on $\mathbb{R}^3 \setminus \{x_0\}$ leads to

$$\begin{aligned} & \int_{\Omega \setminus \Omega^{\infty}} \langle \sigma^{\infty} \nabla u^{\infty}, \nabla \varphi \rangle dV - \int_{\partial \Omega} \langle \sigma^{\infty} \nabla u^{\infty}, \eta \rangle \varphi dS \\ &= \int_{\partial(\Omega \setminus \Omega^{\infty})} \langle \sigma^{\infty} \nabla u^{\infty}, \eta_{\Omega \setminus \Omega^{\infty}} \rangle \varphi dS - \underbrace{\int_{\Omega \setminus \Omega^{\infty}} \text{div}(\sigma^{\infty} \nabla u^{\infty}) \varphi dV}_{=0} - \int_{\partial \Omega} \langle \sigma^{\infty} \nabla u^{\infty}, \eta \rangle \varphi dS \\ &= \int_{\partial \Omega \setminus \partial \Omega_{\text{ext}}^{\infty}} \langle \sigma^{\infty} \nabla u^{\infty}, \eta_{\Omega \setminus \Omega^{\infty}} \rangle \varphi dS + \int_{\partial \Omega_{\text{int}}^{\infty}} \langle \sigma^{\infty} \nabla u^{\infty}, \eta_{\Omega \setminus \Omega^{\infty}} \rangle \varphi dS \\ & \quad - \int_{\partial \Omega \setminus \partial \Omega_{\text{ext}}^{\infty}} \langle \sigma^{\infty} \nabla u^{\infty}, \eta \rangle \varphi dS - \int_{\partial \Omega_{\text{ext}}^{\infty}} \langle \sigma^{\infty} \nabla u^{\infty}, \eta \rangle \varphi dS \\ &= - \int_{\partial \Omega^{\infty}} \langle \sigma^{\infty} \nabla u^{\infty}, \eta_{\Omega^{\infty}} \rangle \varphi dS. \end{aligned}$$

Together with equation (2.40), we thus arrive at

$$\begin{aligned} & \int_{\Omega} \langle \sigma \nabla u^c, \nabla \varphi \rangle dV - \int_{\partial \Omega} \langle \sigma \nabla u, \eta \rangle \varphi dS \\ &= - \int_{\Omega^{\infty}} \langle \sigma^c \nabla u^{\infty}, \nabla \varphi \rangle dV - \int_{\tilde{\Omega}} \langle \sigma \nabla \chi u^{\infty}, \nabla \varphi \rangle dV - \int_{\partial \Omega^{\infty}} \langle \sigma^{\infty} \nabla u^{\infty}, \eta \rangle \varphi dS, \quad (2.41) \end{aligned}$$

where in a slight abuse of notation we now also write η for the outer normal of the patch Ω^{∞} .

We are now done with the right-hand side. For the PEM approach, we have $\langle \sigma \nabla u, \eta \rangle = 0$ on $\partial\Omega$ by definition 2.1.1, and thus also $\int_{\partial\Omega} \langle \sigma \nabla u, \eta \rangle \varphi dS = 0$. In the CEM case, we can instead use definition 2.1.2 to compute for an arbitrary vector $c = (c_1, \dots, c_L) \in \mathbb{R}^L$ that

$$\begin{aligned} \int_{\partial\Omega} \langle \sigma \nabla u, \eta \rangle \varphi dS &= \sum_{l=1}^L \int_{e_l} \langle \sigma \nabla u, \eta \rangle \varphi dS = \sum_{l=1}^L \int_{e_l} \langle \sigma \nabla u, \eta \rangle (\varphi - c_l) dS \\ &= \sum_{l=1}^L \frac{1}{z_l} \cdot \int_{e_l} (U_l - u)(\varphi - c_l) dS \\ &= \sum_{l=1}^L \frac{1}{z_l} \cdot \int_{e_l} (U_l - u^c)(\varphi - c_l) dS - \sum_{l=1}^L \frac{1}{z_l} \cdot \int_{e_l} \chi u^\infty (\varphi - c_l) dS, \end{aligned} \quad (2.42)$$

which when inserted into (2.41) yields

$$\begin{aligned} \int_{\Omega} \langle \sigma \nabla u^c, \nabla \varphi \rangle dV + \sum_{l=1}^L \frac{1}{z_l} \cdot \int_{e_l} (u^c - U_l)(\varphi - c_l) dS \\ = - \int_{\Omega^\infty} \langle \sigma^c \nabla u^\infty, \nabla \varphi \rangle dV - \int_{\tilde{\Omega}} \langle \sigma \nabla \chi u^\infty, \nabla \varphi \rangle dV - \int_{\partial\Omega^\infty} \langle \sigma^\infty \nabla u^\infty, \eta \rangle \varphi dS \\ - \sum_{l=1}^L \frac{1}{z_l} \cdot \int_{e_l} \chi u^\infty (\varphi - c_l) dS. \end{aligned}$$

This finishes the derivation of what we call the *local subtraction approach*.

Definition 2.5.12 (Local subtraction approach). Assume we have a current dipole at position $x_0 \in \Omega$ with moment $M \in \mathbb{R}^3$. Let $V \subset H^1(\Omega)$ be a closed subspace. Let Ω^∞ denote a patch region, $\tilde{\Omega}$ a transition region, and χ a cutoff function. Let a_{PEM} and a_{CEM} be defined as in (2.12) and (2.13). Then, define

$$l_{\text{LS}}^{\text{PEM}} : H^1(\Omega) \rightarrow \mathbb{R}; v \mapsto - \int_{\Omega^\infty} \langle \sigma^c \nabla u^\infty, \nabla v \rangle dV - \int_{\partial\Omega^\infty} \langle \sigma^\infty \nabla u^\infty, \eta \rangle v dS - \int_{\tilde{\Omega}} \langle \sigma \nabla \chi u^\infty, \nabla v \rangle dV,$$

where η denotes the unit outer normal of Ω^∞ . Furthermore, define

$$l_{\text{LS}}^{\text{CEM}} : H^1(\Omega) \oplus \mathbb{R}^L \rightarrow \mathbb{R}; (v, V) \mapsto l_{\text{LS}}^{\text{PEM}}(v) - \sum_{l=1}^L \frac{1}{z_l} \cdot \int_{e_l} \chi u^\infty (v - V_l) dS.$$

Then, the PEM local subtraction approach is given by the problem of finding $u^c \in V$ such that

$$a_{\text{PEM}}(u^c, v) = l_{\text{LS}}^{\text{PEM}}(v)$$

for all $v \in V$. Furthermore, the CEM local subtraction approach is given by the problem of finding $(u^c, U) \in V \oplus \mathbb{R}^L$ such that

$$a_{\text{CEM}}((u^c, U), (v, V)) = l_{\text{LS}}^{\text{CEM}}(v, V)$$

for all $(v, V) \in V \oplus \mathbb{R}^L$. In the PEM case, the final head potential is then given by $u = u^c + \chi \cdot u^\infty$, while in the CEM case, the final head potential is given by $(u^c + \chi u^\infty, U)$.

Note that the choice $\Omega^\infty = \Omega$, $\tilde{\Omega} = \emptyset$, and $\chi = 1$ reproduces the classical subtraction approach as defined in 2.5.10. The classical subtraction approach is thus a special case of the local subtraction approach. Furthermore, note that the functionals $l_{\text{LS}}^{\text{PEM}}$ and $l_{\text{LS}}^{\text{CEM}}$ are non-zero only for those test functions φ that are non-zero on the region where $\chi \neq 0$. If we thus choose this region to be small, only a few test basis functions will evaluate to a non-zero value. We thus see that, for a χ with a small support, the local subtraction approach produces a sparse right-hand side, in contrast to the classical subtraction approach.

Additionally, note that by assumption $\sigma^c = 0$ on $B_\epsilon(x_0)$ for some $\epsilon > 0$, where $B_\epsilon(x_0)$ denotes the balls of radius ϵ around x_0 . By theorem 2.6.6, we have

$$u^\infty(x) = \frac{1}{4\pi\sqrt{\det(\sigma^\infty)}} \frac{\langle M, (\sigma^\infty)^{-1}(x - x_0) \rangle}{\langle (\sigma^\infty)^{-1}(x - x_0), x - x_0 \rangle^{\frac{3}{2}}}.$$

A straightforward computation then yields

$$\begin{aligned} & \nabla u^\infty(x) \\ &= \frac{1}{4\pi\sqrt{\det(\sigma^\infty)}} \cdot \frac{(\sigma^\infty)^{-1} \cdot M}{\|(\sigma^\infty)^{-\frac{1}{2}}(x - x_0)\|^3} - \frac{3}{4\pi\sqrt{\det(\sigma^\infty)}} \cdot \frac{\langle M, (\sigma^\infty)^{-1}(x - x_0) \rangle}{\|(\sigma^\infty)^{-\frac{1}{2}}(x - x_0)\|^5} \cdot (\sigma^\infty)^{-1} \cdot (x - x_0). \end{aligned}$$

We thus see that there is a constant $C > 0$ such that on $\mathbb{R}^3 \setminus \{x_0\}$ we have

$$|u^\infty(x)| \leq C \cdot \frac{1}{\|x - x_0\|^2} \quad (2.43)$$

and

$$\|\nabla u^\infty(x)\| \leq C \cdot \frac{1}{\|x - x_0\|^3}. \quad (2.44)$$

This in particular implies that u^∞ and ∇u^∞ are bounded and smooth on $\mathbb{R}^3 \setminus B_\epsilon(x_0)$, and thus square integrable on bounded sets disjoint from $B_\epsilon(x_0)$. In combination with the trace theorem this implies that $l_{\text{LS}}^{\text{PEM}}$ is a well-defined continuous linear functional $H^1(\Omega)$, and that $l_{\text{LS}}^{\text{CEM}}$ is well-defined, linear, and continuous on $H^1(\Omega) \oplus \mathbb{R}^L$.

We now want to discuss the existence of solutions for the local subtraction approach.

Theorem 2.5.13. *The local subtraction approach as defined in 2.5.12 has a solution for an arbitrary choice of V , and in particular for $V = H^1(\Omega)$, both in the PEM as well as in the CEM case. This solution is unique up to an additive constant, i.e., if $u_0^{c, \text{PEM}}$ is a solution of the PEM problem, and $(u_0^{c, \text{CEM}}, U)$ is a solution of the CEM problem, then the set of all solutions of the PEM problem is given by*

$$u_0^{c, \text{PEM}} + \mathbb{R} \cdot 1,$$

and the set of all solutions of the CEM problem is given by

$$(u_0^{c, \text{CEM}}, U) + \mathbb{R} \cdot (1, 1).$$

Proof. Using corollary 2.4.6 and the strategy from the proof of theorem 2.4.12, the statement of the theorem follows if we can show $l_{\text{LS}}^{\text{PEM}}(1) = 0$ and $l_{\text{LS}}^{\text{CEM}}(1, 1) = 0$. Looking at the definitions of these functionals from 2.5.12, both of these statements reduce to showing that

$$\int_{\partial\Omega} \langle \sigma^\infty \nabla u^\infty, \eta \rangle dS = 0.$$

We derived a proof of this statement that is published in [91]. In the following, we will reproduce this proof as is, with small adjustments to avoid redundancy.

Let $R > 0$ be large enough so that $\overline{\Omega^\infty}$ is contained in the interior of the ball of radius R around x_0 , called $B_R(x_0)$. Let $W = B_R(x_0) \setminus \Omega^\infty$. Since $\operatorname{div}(\sigma^\infty \nabla u^\infty) = 0$ on $\mathbb{R}^3 \setminus \{x_0\}$ the divergence theorem then implies

$$0 = \int_W \operatorname{div}(\sigma^\infty \nabla u^\infty) dV = \int_{\partial W} \langle \sigma^\infty \nabla u^\infty, \eta \rangle dS,$$

where η denotes the outer normal of W . Since we have $\partial W = S_R(x_0) \cup \partial\Omega^\infty$, where $S_R(x_0) = \partial B_R(x_0)$, and the outer normal of W on $\partial\Omega^\infty$ is the negative of the outer normal of Ω^∞ , this implies

$$\int_{\partial\Omega^\infty} \langle \sigma^\infty \nabla u^\infty, \eta \rangle dS = \int_{S_R(x_0)} \langle \sigma^\infty \nabla u^\infty, \eta \rangle dS,$$

where η again denotes the corresponding unit outer normals. Equation (2.44) now yields that

$$\int_{S_R(x_0)} \langle \sigma^\infty \nabla u^\infty, \eta \rangle dS \in \mathcal{O}\left(\frac{1}{R}\right),$$

and hence

$$\int_{\partial\Omega^\infty} \langle \sigma^\infty \nabla u^\infty, \eta \rangle dS = \int_{S_R(x_0)} \langle \sigma^\infty \nabla u^\infty, \eta \rangle dS \xrightarrow{R \rightarrow \infty} 0.$$

This is only possible if

$$\int_{\partial\Omega^\infty} \langle \sigma^\infty \nabla u^\infty, \eta \rangle dS = 0.$$

□

We want to note that this theorem can also be derived in other ways. One example is given in [239], where the authors derive the same statement by exploiting the symmetry of u^∞ around the source position x_0 .

Furthermore, the argument above explicitly allows the choice $V = H^1(\Omega)$. Thus, while theorem 2.5.13 can be used to show that the FEM Galerkin approximation of the local subtraction approach has a solution, the theorem can also be used to show that the continuous EEG forward problem has a solution that is unique up to an additive constant. We want to emphasize that this strategy only works under the assumption that σ is constant in a small environment of the source position. If this were not the case, the functional $l_{\text{LS}}^{\text{PEM}}$ would not be well-defined, as ∇u^∞ behaves like $\frac{1}{\|x-x_0\|^3}$ around the source position x_0 , and we hence cannot expect $\langle \sigma^\infty \nabla u^\infty, \nabla \varphi \rangle$ to be integrable for $\varphi \in H^1(\Omega)$, let alone define a $H^1(\Omega)$ -continuous functional. It is, however, possible to show the existence of a solution to the continuous problem also under much weaker assumptions. In [206] it is e.g. shown that assuming that the components of σ are in $L^\infty(\Omega)$ and are locally Lipschitz continuous around the source position x_0 is already sufficient to derive the existence and uniqueness up to a constant of a solution $u \in L^q(\Omega)$, $1 \leq q < \frac{3}{2}$, of the continuous EEG forward problem for a point dipole. Note that $1 \leq q < \frac{3}{2}$ is exactly the degree of integrability one would expect from comparison with u^∞ , based on (2.43).

We have thus derived the local subtraction approach and shown that it produces solutions that are unique up to an additive constant. But before continuing, we want to address two consistency questions. First of all, the local subtraction approach is supposed to describe the electric potential

in the head. This electric potential does not depend on the choice of cutoff function, patch, and transition region. Thus, ideally, while the correction potential u^c from definition 2.5.12 depends on these choices, the potential $u^c + \chi u^\infty$ should be the same for each choice, at least if we employ a sufficiently large test function space. We shall see that this is indeed the case. Secondly, if one compares definition 2.5.13 to the works of Beltrachini [17], where they derive a CEM version for the classical subtraction approach, one sees that the two formulations look different at first sight. We will prove that they are, in fact, equivalent, both for the continuous formulation as well as for the discrete approximations.

Theorem 2.5.14. *Let $\Omega_1^\infty, \tilde{\Omega}_1, \chi_1$ and $\Omega_2^\infty, \tilde{\Omega}_2, \chi_2$ be two choices of patch, transition region, and cutoff function. For $V = H^1(\Omega)$ in definition 2.5.12, let $u_1^{c,PEM}$ and $u_2^{c,PEM}$ denote the corresponding weak solutions of the PEM local subtraction approach. We then have*

$$u_1^{c,PEM} + \chi_1 \cdot u^\infty + \mathbb{R} \cdot 1 = u_2^{c,PEM} + \chi_2 \cdot u^\infty + \mathbb{R} \cdot 1.$$

Furthermore, if $(u_1^{c,CEM}, U_1)$ and $(u_2^{c,CEM}, U_2)$ denote the corresponding CEM local subtraction solutions, we have

$$(u_1^{c,CEM} + \chi_1 \cdot u^\infty, U_1) + \mathbb{R} \cdot (1, 1) = (u_2^{c,CEM} + \chi_2 \cdot u^\infty, U_2) + \mathbb{R} \cdot (1, 1).$$

Thus, in both the PEM and the CEM case, the total potentials defined by the local subtraction approach are independent of the choice of the patch, transition region, and cutoff function.

Proof. For the PEM case, we published a proof of this theorem in [91]. We will first reproduce this proof here, and then show how to extend it to the CEM case. First of all, note that it suffices to show that each local subtraction approach is equivalent to the classical subtraction approach, and we can WLOG assume that $\Omega_2^\infty = \Omega$, $\tilde{\Omega}_2 = \emptyset$, and $\chi_2 = 1$.

Now note that

$$u_1^{c,PEM} - u_2^{c,PEM} + (\chi_1 - 1) \cdot u^\infty \in H^1(\Omega),$$

since $\chi_1 - 1$ vanishes on an environment of x_0 . Then we have for $v \in H^1(\Omega)$ that

$$\begin{aligned} a_{PEM}(u_1^{c,PEM} - u_2^{c,PEM} + (\chi_1 - 1) \cdot u^\infty, v) \\ = a_{PEM}(u_1^{c,PEM}, v) - a_{PEM}(u_2^{c,PEM}, v) + \int_{\Omega} \langle \sigma \nabla (\chi_1 - 1) u^\infty, \nabla v \rangle dV. \end{aligned} \quad (2.45)$$

Since $u_1^{c,PEM}$ and $u_2^{c,PEM}$ are solutions to their respective weak problems, we get

$$\begin{aligned} a_{PEM}(u_1^{c,PEM}, v) - a_{PEM}(u_2^{c,PEM}, v) \\ = - \int_{\tilde{\Omega}_1} \langle \sigma \nabla \chi_1 \cdot u^\infty, \nabla v \rangle dV - \int_{\partial \Omega_1^\infty} \langle \sigma^\infty \nabla u^\infty, \eta \rangle v dS - \int_{\Omega_1^\infty} \langle \sigma^c \nabla u^\infty, \nabla v \rangle dV \\ + \int_{\Omega} \langle \sigma^c \nabla u^\infty, \nabla v \rangle dV + \int_{\partial \Omega} \langle \sigma^\infty \nabla u^\infty, \eta \rangle v dS \\ = \left(\int_{\partial \Omega} \langle \sigma^\infty \nabla u^\infty, \eta \rangle v dS - \int_{\partial \Omega_1^\infty} \langle \sigma^\infty \nabla u^\infty, \eta \rangle v dS \right) + \int_{\Omega \setminus \Omega_1^\infty} \langle \sigma^c \nabla u^\infty, \nabla v \rangle dV \\ - \int_{\tilde{\Omega}_1} \langle \sigma \nabla \chi_1 \cdot u^\infty, \nabla v \rangle dV. \end{aligned}$$

Now, just as in the derivation of the local subtraction approach, we have

$$\begin{aligned} \int_{\partial\Omega} \langle \sigma^\infty \nabla u^\infty, \eta \rangle v \, dS - \int_{\partial\Omega_1^\infty} \langle \sigma^\infty \nabla u^\infty, \eta \rangle v \, dS &= \int_{\partial(\Omega \setminus \Omega_1^\infty)} \langle \sigma^\infty \nabla u^\infty, \eta \rangle v \, dS \\ &= \int_{\Omega \setminus \Omega_1^\infty} \underbrace{\operatorname{div}(\sigma^\infty \nabla u^\infty)}_{=0} v + \langle \sigma^\infty \nabla u^\infty, \nabla v \rangle \, dV = \int_{\Omega \setminus \Omega_1^\infty} \langle \sigma^\infty \nabla u^\infty, \nabla v \rangle \, dV \end{aligned}$$

Inserting this into the preceding equation yields

$$\begin{aligned} a_{\text{PEM}}(u_1^{\text{c,PEM}}, v) - a_{\text{PEM}}(u_2^{\text{c,PEM}}, v) \\ = \int_{\Omega \setminus \Omega_1^\infty} \langle \sigma^\infty \nabla u^\infty, \nabla v \rangle \, dV + \int_{\Omega \setminus \Omega_1^\infty} \langle \sigma^c \nabla u^\infty, \nabla v \rangle \, dV - \int_{\tilde{\Omega}_1} \langle \sigma \nabla \chi_1 \cdot u^\infty, \nabla v \rangle \, dV \\ = \int_{\Omega \setminus \Omega_1^\infty} \langle \sigma \nabla (1 - \chi_1) u^\infty, \nabla v \rangle \, dV. \end{aligned}$$

Inserting this into (2.45) and noting that $\chi_1 - 1 = 0$ on Ω_1^∞ , we see that

$$a_{\text{PEM}}(u_1^{\text{c,PEM}} - u_2^{\text{c,PEM}} + (\chi_1 - 1) \cdot u^\infty, v) = 0.$$

This implies that $(u_1^{\text{c,PEM}} + \chi_1 \cdot u^\infty) - (u_2^{\text{c,PEM}} + u^\infty) = c$ for some constant $c \in \mathbb{R}$, which shows the claim of the theorem in the PEM case.

For the CEM case, we similarly have

$$(u_1^{\text{c,CEM}} + \chi_1 \cdot u^\infty, U_1) - (u_2^{\text{c,CEM}} + u^\infty, U_2) = (u_1^{\text{c,CEM}} - u_2^{\text{c,CEM}} + (\chi_1 - 1) \cdot u^\infty, U_1 - U_2) \in H^1(\Omega) \oplus \mathbb{R}^L,$$

and thus for $(v, V) \in H^1(\Omega) \oplus \mathbb{R}^L$

$$\begin{aligned} a_{\text{CEM}}((u_1^{\text{c,CEM}} - u_2^{\text{c,CEM}} + (\chi_1 - 1) \cdot u^\infty, U_1 - U_2), (v, V)) \\ = a_{\text{CEM}}((u_1^{\text{c,CEM}}, U_1), (v, V)) - a_{\text{CEM}}((u_2^{\text{c,CEM}}, U_2), (v, V)) + a_{\text{CEM}}(((\chi_1 - 1) \cdot u^\infty, 0), (v, V)) \end{aligned}$$

For the third term, we have, by definition of a_{CEM} , that

$$a_{\text{CEM}}(((\chi_1 - 1) \cdot u^\infty, 0), (v, V)) = \int_{\Omega} \langle \sigma \nabla (\chi_1 - 1) \cdot u^\infty, \nabla v \rangle \, dV + \sum_{l=1}^L \frac{1}{z_l} \int_{e_l} (\chi_1 - 1) \cdot u^\infty \cdot (v - V_l) \, dS.$$

Since $(u_1^{\text{c,CEM}}, U_1)$ and $(u_2^{\text{c,CEM}}, U_2)$ are solutions to their corresponding weak formulations, we furthermore get for the first two terms that

$$\begin{aligned} a_{\text{CEM}}((u_1^{\text{c,CEM}}, U_1), (v, V)) - a_{\text{CEM}}((u_2^{\text{c,CEM}}, U_2), (v, V)) \\ = l_{\text{LS},1}^{\text{PEM}}(v) - \sum_{l=1}^L \frac{1}{z_l} \int_{e_l} \chi_1 \cdot u^\infty (v - V_l) \, dS - l_{\text{LS},2}^{\text{PEM}}(v) + \sum_{l=1}^L \frac{1}{z_l} \int_{e_l} u^\infty (v - V_l) \, dS, \end{aligned}$$

where $l_{\text{LS},i}^{\text{PEM}}$ denotes the local subtraction right-hand side functional for the patch, transition, and cutoff choice Ω_i^∞ , $\tilde{\Omega}_i$, and χ_i . But now note that the terms corresponding to integrals over the electrode interfaces e_l exactly cancel, and we have

$$\begin{aligned} a_{\text{CEM}}((u_1^{\text{c,CEM}} - u_2^{\text{c,CEM}} + (\chi_1 - 1) \cdot u^\infty, U_1 - U_2), (v, V)) \\ = l_{\text{LS},1}^{\text{PEM}}(v) - l_{\text{LS},2}^{\text{PEM}}(v) + \int_{\Omega} \langle \sigma \nabla (\chi_1 - 1) \cdot u^\infty, \nabla v \rangle \, dV. \end{aligned}$$

The right-hand side of this expression is now exactly the same term that arose during the PEM part of this proof, and, by the same argument as in the PEM part, it is equal to zero. We have thus shown that

$$a_{\text{CEM}}((u_1^{c,\text{CEM}} - u_2^{c,\text{CEM}} + (\chi_1 - 1) \cdot u^\infty, U_1 - U_2), (v, V)) = 0$$

for all $(v, V) \in H^1(\Omega) \oplus \mathbb{R}^L$. By lemma 2.4.3 it follows that

$$(u_1^{c,\text{CEM}} + \chi_1 \cdot u^\infty, U_1) - (u_2^{c,\text{CEM}} + u^\infty, U_2) = c \cdot (1, 1)$$

for some $c \in \mathbb{R}$, which finishes the proof. \square

We have thus seen that for the continuous weak formulation, the choice of patch, transition region, and cutoff function has no impact on the head potential. This is, however, not true for the discrete Galerkin approximations. In section 2.8.2, we will investigate the impact of this choice on the accuracy of the FEM simulation.

If we look at the definition of the CEM local subtraction approach in 2.5.12, one might get the impression that the treatment of the volume part of the potential and the electrode part of the potential is somewhat asymmetric, as we subtract the infinity potential from u , but not from U . Alternatively, one might choose some values U_l^∞ , $1 \leq l \leq L$, and try to solve for $U_l^c := U - U_l^\infty$. In the context of the classical subtraction approach, this is e.g. done in [17], where the author sets $U_l^\infty := \frac{\int_{e_l} u^\infty dS}{\text{area}(e_l)}$, and solves for a solution (u^c, U^c) fulfilling

$$\begin{aligned} a_{\text{CEM}}((u^c, U^c), (v, V)) \\ = - \int_{\Omega} \langle \sigma^c \nabla u^c, \nabla v \rangle dV - \int_{\partial\Omega} \langle \sigma^\infty \nabla u^\infty, \eta \rangle v dS - \sum_{l=1}^L \frac{1}{z_l} \int_{e_l} (u^\infty - U_l^\infty) \cdot (v - V_l) dS. \end{aligned}$$

The final head potential is then given by $(u^c, U^c) + (u^\infty, U^\infty)$, where $U^\infty = (U_1^\infty, \dots, U_L^\infty)$. We want to show that this is exactly equivalent to our definition of the subtraction CEM approach from 2.5.12, both in the continuous as well as in the discrete case. This is a special case of the following result.

Theorem 2.5.15. *Let $V \subset H^1(\Omega)$ be a closed subspace with $(1, 1) \in V$, and let Ω^∞ , $\tilde{\Omega}$, and χ denote an arbitrary choice of patch, transition region, and cutoff function. Let $(u^c, U) \in V \oplus \mathbb{R}^L$ denote a solution of the CEM local subtraction approach, as defined in 2.5.12. Now, let $c = (c_1, \dots, c_L) \in \mathbb{R}^L$ denote an arbitrary vector. Define a functional*

$$l : H^1(\Omega) \oplus \mathbb{R}^L \rightarrow \mathbb{R}; (v, V) \mapsto l_{LS}^{\text{PEM}}(v) - \sum_{l=1}^L \frac{1}{z_l} \int_{e_l} (\chi \cdot u^\infty - c_l) \cdot (v - V_l) dS.$$

Then, l is a continuous linear functional with $l(1, 1) = 0$, and the problem

$$\text{Find } (\tilde{u}^c, U^C) \in V \oplus \mathbb{R}^L \text{ such that } a_{\text{CEM}}((\tilde{u}^c, U^C), (v, V)) = l(v, V) \text{ for all } (v, V) \in V \oplus \mathbb{R}^L$$

has a solution that is unique up to adding a multiple of $(1, 1)$. Furthermore, we have

$$(u^c + \chi \cdot u^\infty, U) + \mathbb{R} \cdot (1, 1) = (\tilde{u}^c + \chi \cdot u^\infty, U^c + c) + \mathbb{R} \cdot (1, 1).$$

Thus, l defines the same solution family as the local subtraction CEM approach from 2.5.12.

Proof. The functional l is linear, continuous, and fulfills $l(1, 1) = 0$ by the same argument as for $l_{\text{LS}}^{\text{CEM}}$. By the same argument as before, this implies that the corresponding weak formulation has a solution that is unique up to adding a multiple of $(1, 1)$. Now, let $(\widehat{u}^c, \widehat{U}) := (\widehat{u}^c, U^c + c) \in V \oplus \mathbb{R}^L$. Then, the statement of the theorem follows if we can show that

$$a_{\text{CEM}}((\widehat{u}^c, \widehat{U}), (v, V)) = l_{\text{LS}}^{\text{CEM}}(v, V).$$

We can now compute

$$\begin{aligned} a_{\text{CEM}}((\widehat{u}^c, \widehat{U}), (v, V)) &= a_{\text{CEM}}((\widehat{u}^c, U^c), (v, V)) + a_{\text{CEM}}((0, c), (v, V)) \\ &= l(v, V) - \sum_{l=1}^L \frac{1}{z_l} \int_{e_l} c_l \cdot (v - V_l) dS \\ &= l_{\text{LS}}^{\text{PEM}}(v) - \sum_{l=1}^L \frac{1}{z_l} \int_{e_l} \chi \cdot u^\infty (v - V_l) dS \\ &= l_{\text{LS}}^{\text{CEM}}(v, V). \end{aligned}$$

□

Note that we could have naturally arrived at the functional l during our derivation of the local subtraction approach by introducing the corresponding constants in equation (2.42). This, for now, finishes our investigations into the theoretical properties of the local subtraction approach. One natural question that is left open at this point is the relation between the CEM and the PEM approaches for the EEG forward problem. In theorem 2.4.8, it is shown that, if the CEM electrode interfaces shrink toward single points, the CEM tDCS solution in the interior of the head converges to the PEM tDCS solution. One might expect that one can get a similar result for the EEG forward problem. This is, in fact, the case. Once we have derived the reciprocity theorems, we will show in theorem 2.9.10 how the relation between the EEG and tDCS problems, in combination with elliptic regularity, can be used to directly transfer the tDCS convergence result to the EEG forward problem for point dipoles.

We have now discussed the current state-of-the-art of EEG forward modeling, and extended the theory with a new approach, which we termed the *local subtraction approach*. This brings us, finally, to the MEG forward problem.

2.5.3 The MEG forward problem

In definition 2.1.3, we derived that the MEG forward problem consists of evaluating the integral

$$B(x) = \frac{\mu_0}{4\pi} \cdot \int_{\Omega} j^{\text{P}} \times \frac{x - y}{\|x - y\|^3} dV(y) - \frac{\mu_0}{4\pi} \cdot \int_{\Omega} \sigma \nabla u \times \frac{x - y}{\|x - y\|^3} dV(y), \quad (2.46)$$

where j^{P} is the neural activity and u is the electric potential in the head generated by the neural activity. Thus, B depends linearly on the neural activity j^{P} , and we can again restrict ourselves to the case of a single point dipole. If $j^{\text{P}} = M \cdot \delta_{x_0}$, we get for the primary magnetic field

$$\begin{aligned} B^{\text{P}}(x) &= \frac{\mu_0}{4\pi} \cdot \int_{\Omega} j^{\text{P}} \times \frac{x - y}{\|x - y\|^3} dV(y) = \frac{\mu_0}{4\pi} \cdot \int_{\Omega} M \times \frac{x - y}{\|x - y\|^3} \cdot \delta_{x_0}(y) dV(y) \\ &= \frac{\mu_0}{4\pi} \cdot M \times \frac{x - x_0}{\|x - x_0\|^3}. \end{aligned}$$

Thus, B^P is easy to compute. The difficult part of the MEG forward problem lies in the evaluation of the integral

$$\int_{\Omega} \sigma \nabla u \times \frac{x-y}{\|x-y\|^3} dV(y).$$

More concretely, MEG devices typically measure the magnetic field in a certain direction, say $v \in \mathbb{R}^3$, and we are interested in evaluating

$$\int_{\Omega} \langle \sigma \nabla u \times \frac{x-y}{\|x-y\|^3}, v \rangle dV(y).$$

In the standard case, we approximate u using a FEM simulation. If one employs one of the direct approaches from section 2.5.1, the integral above poses no significant problem, as the result of the FEM simulation is given in the form $u_h = \sum_{i=1}^N c_i \cdot \varphi_i$, where $\varphi_1, \dots, \varphi_N$ is a basis of the FEM trial space. This leads to

$$\int_{\Omega} \langle \sigma \nabla u_h \times \frac{x-y}{\|x-y\|^3}, v \rangle dV(y) = \sum_{i=1}^N \left(\int_{\Omega} \langle \sigma \nabla \varphi_i \times \frac{x-y}{\|x-y\|^3}, v \rangle dV(y) \right) \cdot c_i.$$

The integrals over the FEM basis functions on the right-hand side can be pre-computed. Then, the MEG forward problem can simply be solved by first solving the EEG forward problem and then computing an inner product.

Definition 2.5.16 (MEG forward problem with direct EEG forward approach). Assume that we want to solve the MEG forward problem for a dipole at position $x_0 \in \Omega$ with moment $M \in \mathbb{R}^3$. If the solution of the EEG forward problem was computed using a direct FEM forward approach, it is given in the form $u = \sum_{i=1}^N u_i \cdot \varphi$, where $\varphi_1, \dots, \varphi_N$ is a basis of the FEM trial space. The MEG forward problem for a coil at position $x \in \mathbb{R}^3 \setminus \Omega$ with direction $v \in \mathbb{R}^3$ is then solved by computing the value

$$\langle B(x), v \rangle = \frac{\mu_0}{4\pi} \langle M \times \frac{x-x_0}{\|x-x_0\|^3}, v \rangle - \frac{\mu_0}{4\pi} \cdot \sum_{i=1}^N \left(\int_{\Omega} \langle \sigma \nabla \varphi_i \times \frac{x-y}{\|x-y\|^3}, v \rangle dV(y) \right) \cdot u_i.$$

If one instead employs a subtraction approach, this strategy runs into difficulties. If, e.g., u_h^c denotes the result of a local subtraction FEM simulation, as defined in 2.5.12, the approximation to the head potential is given by $u = u^c + \chi \cdot u^\infty$. But close to the source position x_0 we have $\chi = 1$, and thus equation (2.46) leads to integrals of the form

$$\int_U \langle \sigma \nabla u^\infty \times \frac{x-y}{\|x-y\|^3}, v \rangle dV, \quad (2.47)$$

where U is an environment of x_0 . Based on (2.44), we expect that close to the source position x_0 the gradient of the homogeneous potential ∇u^∞ behaves like $\frac{1}{\|x-x_0\|^3}$. Explicitly writing out the integrand for a concrete example, e.g. for $M = e_1$, $\sigma^\infty = E_3$, $x_0 = 0$, and $v = e_1$, it is in fact not difficult to see that $\langle \sigma \nabla u^\infty \times \frac{x-y}{\|x-y\|^3}, v \rangle$ is in general not integrable on a small environment of x_0 , and it is hence not even clear what the expression (2.47) is supposed to mean.

If the conductivity σ^∞ in the environment of the source position is anisotropic, there is, to the best of our knowledge, no known strategy to solve this problem. Instead, researchers have typically used ad hoc “fixes”, such as simply ignoring the mesh element containing the source position in the evaluation of the integral, or replacing u^∞ by an interpolation from the trial function space [152].

If the conductivity σ^∞ is instead isotropic, one can proceed in an “analytic” way, and we make the assumption that σ^∞ is isotropic for the remainder of this section. The idea on how to treat the problem of the integral not being defined goes back to Geselowitz [71], and is utilized both in the BEM theory [83] as well as in the FEM subtraction approach [151], and is also what we used in our own publication on the local subtraction approach [91]. Plainly speaking, the idea consists of simply ignoring that the integral is not defined, applying the divergence theorem anyway, and thus rewriting the undefined volume integral as a well-defined surface integral. On first reading, this might cause some stomach pain, but the fact that this approach was, e.g., published in [83] in 1993, and has since been used in numerous studies where it was found to lead to numerical approaches that agree with measurements, suggests that the approach indeed leads to the “correct” formulation.

We refer to our publication [91] for the concrete details on this derivation. Here, we instead want to show a different derivation, which starts directly from Maxwell’s equations, and hence avoids the detour over the Biot-Savart law and some of the accompanying technical difficulties mentioned above.

First, the quasi-static approximations of Maxwell’s equations (2.3) and (2.4) stipulate that the magnetic field B due to a dipole at position x_0 with moment M is given by

$$\begin{aligned}\operatorname{div}(B) &= 0 \\ \operatorname{rot}(B) &= \mu_0 \cdot (M \cdot \delta_{x_0} - \sigma \nabla u)\end{aligned}$$

Now remember that in theorem 2.2.4 we have shown that, in an unbounded, homogeneous, and isotropic volume conductor, the total magnetic field is already given by the primary magnetic field. If we thus define

$$B^\infty(x) = \frac{\mu_0}{4\pi} \cdot M \times \frac{x - x_0}{\|x - x_0\|^3},$$

it follows that B^∞ fulfills

$$\operatorname{div}(B^\infty) = 0 \tag{2.48}$$

$$\operatorname{rot}(B^\infty) = \mu_0 \cdot (M \cdot \delta_{x_0} - \sigma^\infty \nabla u^\infty) \tag{2.49}$$

Note that it is also straightforward to verify these last two equations with a direct computation, since

$$B^\infty(x) = \operatorname{rot} \left(\frac{\mu_0}{4\pi} \cdot \frac{M}{\|x - x_0\|} \right)$$

implies that

$$\operatorname{div}(B^\infty) = \operatorname{div}(\operatorname{rot} \left(\frac{\mu_0}{4\pi} \cdot \frac{M}{\|x - x_0\|} \right)) = 0,$$

and we have shown equation (2.48). Furthermore, $\operatorname{rot}(\operatorname{rot}(F)) = \nabla(\operatorname{div}(F)) - \Delta(F)$ for a vector field F implies that

$$\operatorname{rot}(B^\infty) = \operatorname{rot}(\operatorname{rot} \left(\frac{\mu_0}{4\pi} \cdot \frac{M}{\|x - x_0\|} \right)) = \nabla \left(\operatorname{div} \left(\frac{\mu_0}{4\pi} \cdot \frac{M}{\|x - x_0\|} \right) \right) - \Delta \left(\frac{\mu_0}{4\pi} \cdot \frac{M}{\|x - x_0\|} \right).$$

Now, we have

$$\operatorname{div} \left(\frac{1}{4\pi} \cdot \frac{M}{\|x - x_0\|} \right) = \frac{1}{4\pi} \langle M, \nabla_x \frac{1}{\|x - x_0\|} \rangle = -\sigma^\infty \cdot \frac{1}{4\pi\sigma^\infty} \cdot \frac{\langle M, x - x_0 \rangle}{\|x - x_0\|^3} = -\sigma^\infty u^\infty,$$

where we have used the formula for the isotropic u^∞ from section 2.6.2. This yields

$$\nabla \left(\operatorname{div} \left(\frac{\mu_0}{4\pi} \cdot \frac{M}{\|x - x_0\|} \right) \right) = -\mu_0 \cdot \sigma^\infty \nabla u^\infty.$$

Furthermore, since the Newton potential is a fundamental solution for the Laplace operator (see e.g. [64], §17, Satz 2), we get

$$-\Delta \left(\frac{\mu_0}{4\pi} \cdot \frac{M}{\|x - x_0\|} \right) = \mu_0 \cdot M \cdot \delta_{x_0},$$

which shows equation (2.49).

We can now compute for the secondary field $B^S := B - B^P = B - B^\infty$ that

$$\begin{aligned} \operatorname{div}(B^S) &= \operatorname{div}(B) - \operatorname{div}(B^\infty) = 0 \\ \operatorname{rot}(B^S) &= \operatorname{rot}(B) - \operatorname{rot}(B^\infty) = \mu_0(\sigma^\infty \nabla u^\infty - \sigma \nabla u) \end{aligned}$$

Now, since σ^∞ is constant and isotropic, we get

$$\begin{aligned} \Delta B^S &= \nabla (\operatorname{div}(B^S)) - \operatorname{rot}(\operatorname{rot}(B^S)) = \mu_0 \operatorname{rot}(\sigma \nabla u) - \mu_0 \sigma^\infty \operatorname{rot}(\nabla u^\infty) \\ &= \mu_0 \operatorname{rot}(\sigma \nabla u) = \mu_0 \operatorname{rot}(\sigma^c \nabla u + \sigma^\infty \nabla u) \\ &= \mu_0 \operatorname{rot}(\sigma^c \nabla u). \end{aligned}$$

We thus see that both the dipole $M \cdot \delta_{x_0}$ itself as well as the singularity of the volume current $\sigma \nabla u$ in the proximity of the dipole position have been eliminated from the problem, as $\sigma^c = 0$ in an environment of the source position. But now we can proceed in the standard way and obtain B^S via convolution with the fundamental solution of the Laplace operator. To this end, let Ω^∞ , $\tilde{\Omega}$, and χ be an arbitrary choice of patch, transition region, and cutoff function, as defined in 2.5.12, and let u^c denote the corresponding local subtraction correction potential. Then, on the head Ω the potential u is given by $u = u^c + \chi \cdot u^\infty$, and on Ω we thus have $\sigma^c \nabla u = \sigma^c \nabla u^c + \sigma^c \nabla \chi u^\infty$, while on the outside of Ω we have $\sigma = 0$ and hence $\sigma^c = -\sigma^\infty$. Furthermore, we assume that ∇u decays at infinity like $\frac{1}{R^3}$. Again using that, for a scalar field φ and a vector field F , we have

$$\operatorname{rot}(\varphi \cdot F) = \varphi \cdot \operatorname{rot}(F) + \nabla \varphi \times F$$

and that for a sufficiently regular bounded domain U we have

$$\int_U \operatorname{rot}(F) dV = \int_{\partial U} \eta \times F dS,$$

we can now compute

$$\begin{aligned} B^S(x) &= -\frac{\mu_0}{4\pi} \int_{\mathbb{R}^3} \frac{\operatorname{rot}(\sigma^c \nabla u)(y)}{\|x - y\|} dy = \frac{\mu_0}{4\pi} \int_{\mathbb{R}^3} \nabla_y \left(\frac{1}{\|x - y\|} \right) \times \sigma^c \nabla u(y) dy \\ &= -\frac{\mu_0}{4\pi} \int_{\mathbb{R}^3} \sigma^c \nabla u \times \frac{x - y}{\|x - y\|^3} dy \\ &= \frac{\mu_0}{4\pi} \int_{\mathbb{R}^3 \setminus \Omega} \sigma^\infty \nabla u \times \frac{x - y}{\|x - y\|^3} dy - \frac{\mu_0}{4\pi} \int_{\Omega} \sigma^c \nabla u \times \frac{x - y}{\|x - y\|^3} dy. \end{aligned}$$

If now $B_\epsilon(x)$ denotes the ball of radius ϵ around x , we can compute

$$\begin{aligned} \int_{\mathbb{R}^3 \setminus \Omega} \sigma^\infty \nabla u \times \frac{x-y}{\|x-y\|^3} dy &= \sigma^\infty \int_{\mathbb{R}^3 \setminus \Omega} \operatorname{rot} \left(u \cdot \frac{x-y}{\|x-y\|^3} \right) dy \\ &= - \int_{\partial\Omega} \sigma^\infty u \cdot \eta \times \frac{x-y}{\|x-y\|^3} dS(y) - \lim_{\epsilon \rightarrow 0} \int_{\partial B_\epsilon(x)} \sigma^\infty u \cdot \eta_{B_\epsilon(x)} \times \frac{x-y}{\|x-y\|^3} dS(y). \end{aligned}$$

But now note that the outer normal of $B_\epsilon(x)$ at y is given by $\frac{y-x}{\|y-x\|}$, and we hence have that $\eta_{B_\epsilon(x)}(y) \times \frac{x-y}{\|x-y\|^3} = 0$, and the second integral on the right-hand side above vanishes. We thus have

$$\int_{\mathbb{R}^3 \setminus \Omega} \sigma^\infty \nabla u \times \frac{x-y}{\|x-y\|^3} dy = - \int_{\partial\Omega} \sigma^\infty u \cdot \eta \times \frac{x-y}{\|x-y\|^3} dS(y).$$

Inserting $u = u^c + \chi \cdot u^\infty$ then yields

$$\begin{aligned} - \int_{\partial\Omega} \sigma^\infty u \cdot \eta \times \frac{x-y}{\|x-y\|^3} dS(y) \\ = - \int_{\partial\Omega} (\sigma^\infty \chi u^\infty) \cdot \eta \times \frac{x-y}{\|x-y\|^3} dS(y) - \int_{\Omega} \sigma^\infty \nabla u^c \times \frac{x-y}{\|x-y\|^3} dy. \end{aligned}$$

Inserting this into the expression for B^S above leads to

$$\begin{aligned} B^S(x) &= - \frac{\mu_0}{4\pi} \int_{\partial\Omega} (\sigma^\infty \chi u^\infty) \cdot \eta \times \frac{x-y}{\|x-y\|^3} dS(y) - \frac{\mu_0}{4\pi} \int_{\Omega} \sigma^c \nabla \chi u^\infty \times \frac{x-y}{\|x-y\|^3} dy \\ &\quad - \frac{\mu_0}{4\pi} \int_{\Omega} \sigma^\infty \nabla u^c \times \frac{x-y}{\|x-y\|^3} dy. \end{aligned}$$

Remembering that on $\Omega \setminus (\widetilde{\Omega} \cup \Omega^\infty)$ we have $\chi = 0$, and that on Ω^∞ we have $\chi = 1$, we furthermore have

$$\frac{\mu_0}{4\pi} \int_{\Omega} \sigma^c \nabla \chi u^\infty \times \frac{x-y}{\|x-y\|^3} dy = \frac{\mu_0}{4\pi} \int_{\widetilde{\Omega}} \sigma^c \nabla \chi u^\infty \times \frac{x-y}{\|x-y\|^3} dy + \frac{\mu_0}{4\pi} \int_{\Omega^\infty} \sigma^c \nabla u^\infty \times \frac{x-y}{\|x-y\|^3} dy.$$

This finishes the derivation, which we now summarize in the following definition.

Definition 2.5.17 (MEG forward problem with local subtraction EEG forward approach). Let Ω^∞ , $\widetilde{\Omega}$, and χ denote a choice of patch, transition region, and cutoff function, and let u^c denote the corresponding local subtraction correction potential, as defined in 2.5.12. If u^c was computed using a FEM simulation, it can be written in the form $u^c = \sum_{i=1}^N u_i^c \cdot \varphi_i$, where $\varphi_1, \dots, \varphi_N$ denotes a basis of the FEM trial space. The MEG forward problem for a coil at position $x \in \mathbb{R}^3 \setminus \Omega$ with direction $v \in \mathbb{R}$ is then solved by computing the value

$$\begin{aligned} \langle B(x), v \rangle &= \frac{\mu_0}{4\pi} \cdot \langle M \times \frac{x-x_0}{\|x-x_0\|^3}, v \rangle - \frac{\mu_0}{4\pi} \cdot \sum_{i=1}^N \left(\int_{\Omega} \langle \sigma^\infty \nabla \varphi_i \times \frac{x-y}{\|x-y\|^3}, v \rangle dV(y) \right) \cdot u_i^c \\ &\quad - \frac{\mu_0}{4\pi} \int_{\widetilde{\Omega}} \langle \sigma^c \nabla \chi u^\infty \times \frac{x-y}{\|x-y\|^3}, v \rangle dy - \frac{\mu_0}{4\pi} \int_{\Omega^\infty} \langle \sigma^c \nabla u^\infty \times \frac{x-y}{\|x-y\|^3}, v \rangle dy \\ &\quad - \frac{\mu_0}{4\pi} \int_{\partial\Omega} \langle (\sigma^\infty \chi u^\infty) \cdot \eta \times \frac{x-y}{\|x-y\|^3}, v \rangle dS(y). \end{aligned}$$

Note that, as in the EEG case, this value can be assembled rapidly for many sources. The primary field contribution is a simple and quickly evaluated analytical expression. The integrals over the transition region $\tilde{\Omega}$ and the patch Ω^∞ are also quickly computed as long as both of these sets are chosen to be small, and the integral over $\partial\Omega$ only needs to be assembled on the part of the boundary where $\chi \neq 0$, which will in most cases be nowhere, i.e. we can almost always skip this integral. The only expression where, at this point in this thesis, it is not yet clear how it can be quickly computed is the expression

$$\sum_{i=1}^N \left(\int_{\Omega} \langle \sigma \nabla \varphi_i \times \frac{x-y}{\|x-y\|^3}, v \rangle dV(y) \right) \cdot u_i^c,$$

and the similar expression for the direct approaches in 2.5.16. In section 2.9.4, we will see how to rapidly evaluate expressions of this form by using *reciprocity laws* (which are sometimes also called *transfer matrices* in the literature).

When comparing definition 2.5.17 with the formulation we have published in [91], one sees that they look different. While above we have derived

$$\begin{aligned} B^S(x) = & -\frac{\mu_0}{4\pi} \int_{\partial\Omega} (\sigma^\infty \chi u^\infty) \cdot \eta \times \frac{x-y}{\|x-y\|^3} dS(y) - \frac{\mu_0}{4\pi} \int_{\tilde{\Omega}} \sigma^c \nabla \chi u^\infty \times \frac{x-y}{\|x-y\|^3} dy \\ & - \frac{\mu_0}{4\pi} \int_{\Omega^\infty} \sigma^c \nabla u^\infty \times \frac{x-y}{\|x-y\|^3} dy - \frac{\mu_0}{4\pi} \int_{\Omega} \sigma \nabla u^c \times \frac{x-y}{\|x-y\|^3} dy, \end{aligned}$$

the derivation in [91] yielded

$$\begin{aligned} B^S(x) = & -\frac{\mu_0}{4\pi} \int_{\partial\Omega^\infty} (\sigma^\infty u^\infty) \cdot \eta \times \frac{x-y}{\|x-y\|^3} dS(y) - \frac{\mu_0}{4\pi} \int_{\tilde{\Omega}} \sigma \nabla \chi u^\infty \times \frac{x-y}{\|x-y\|^3} dy \\ & - \frac{\mu_0}{4\pi} \int_{\Omega^\infty} \sigma^c \nabla u^\infty \times \frac{x-y}{\|x-y\|^3} dy - \frac{\mu_0}{4\pi} \int_{\Omega} \sigma \nabla u^c \times \frac{x-y}{\|x-y\|^3} dy. \end{aligned}$$

Note the difference in the first two summands on the right-hand side. Before showing that these two formulations are indeed equivalent, notice that the first version is computationally more efficient than the version in the published paper. Concretely, since the patch will almost never touch the domain boundary, the surface integral over $\partial\Omega$ in the first expression will almost always be zero, while in the second expression the integral over $\partial\Omega^\infty$ always needs to be assembled. Furthermore, the integral over the transition region $\tilde{\Omega}$ in the first formulation contains a factor of σ^c , which implies that on transition elements where $\sigma = \sigma^\infty$, i.e. typically the transition elements in the brain compartment, we do not need to compute an integral and can hence skip over them in the assembly loop. In the second formulation, these elements in general give a non-zero contribution to the integral.

When benchmarking these two strategies using the realistic head model from [91] on an AMD Ryzen Threadripper 3960X processor with 24 cores, these two effects lead to a speedup of ca. 40% when using the formulation from definition 2.5.17 instead of the one published in [91].

Lemma 2.5.18. *The local subtraction MEG formulation from definition 2.5.17 is equivalent to the formulation published in [91].*

Proof. After cancelling the terms that are the same in both formulations and multiplying by

$-\frac{4\pi}{\mu_0}$, we need to show that

$$\begin{aligned} \int_{\partial\Omega} (\sigma^\infty \chi u^\infty) \cdot \eta \times \frac{x-y}{\|x-y\|^3} dS(y) + \int_{\tilde{\Omega}} \sigma^c \nabla \chi u^\infty \times \frac{x-y}{\|x-y\|^3} dy \\ = \int_{\partial\Omega^\infty} (\sigma^\infty u^\infty) \cdot \eta \times \frac{x-y}{\|x-y\|^3} dS(y) + \int_{\tilde{\Omega}} \sigma \nabla \chi u^\infty \times \frac{x-y}{\|x-y\|^3} dy. \end{aligned}$$

This equation is now equivalent to

$$\begin{aligned} \int_{\partial\Omega} (\sigma^\infty \chi u^\infty) \cdot \eta \times \frac{x-y}{\|x-y\|^3} dS(y) - \int_{\partial\Omega^\infty} (\sigma^\infty u^\infty) \cdot \eta \times \frac{x-y}{\|x-y\|^3} dS(y) \\ = \int_{\tilde{\Omega}} \sigma^\infty \nabla \chi u^\infty \times \frac{x-y}{\|x-y\|^3} dy \quad (2.50) \end{aligned}$$

But now we can mimic the derivation of the EEG local subtraction approach. Concretely, let $\partial\Omega_{\text{ext}}^\infty = \partial\Omega^\infty \cap \partial\Omega$, and let $\partial\Omega_{\text{int}}^\infty = \partial\Omega^\infty \setminus \partial\Omega_{\text{ext}}^\infty$. Then, since by construction we have $\chi = 1$ on $\partial\Omega^\infty$, the parts of the integral over $\partial\Omega_{\text{ext}}^\infty$ on the left-hand side of the above integral cancel each other, and we get

$$\begin{aligned} \int_{\partial\Omega} (\sigma^\infty \chi u^\infty) \cdot \eta \times \frac{x-y}{\|x-y\|^3} dS(y) - \int_{\partial\Omega^\infty} (\sigma^\infty u^\infty) \cdot \eta \times \frac{x-y}{\|x-y\|^3} dS(y) \\ = \int_{\partial\Omega} (\sigma^\infty \chi u^\infty) \cdot \eta \times \frac{x-y}{\|x-y\|^3} dS(y) - \int_{\partial\Omega^\infty} (\sigma^\infty \chi u^\infty) \cdot \eta \times \frac{x-y}{\|x-y\|^3} dS(y) \\ = \int_{\partial\Omega \setminus \partial\Omega_{\text{ext}}^\infty} (\sigma^\infty \chi u^\infty) \cdot \eta \times \frac{x-y}{\|x-y\|^3} dS(y) + \int_{\partial\Omega_{\text{int}}^\infty} (\sigma^\infty \chi u^\infty) \cdot \eta_{\Omega \setminus \Omega^\infty} \times \frac{x-y}{\|x-y\|^3} dS(y) \\ = \int_{\partial(\Omega \setminus \Omega^\infty)} (\sigma^\infty \chi u^\infty) \cdot \eta_{\Omega \setminus \Omega^\infty} \times \frac{x-y}{\|x-y\|^3} dS(y) = \int_{\Omega \setminus \Omega^\infty} \text{rot} \left(\sigma^\infty \chi u^\infty \cdot \frac{x-y}{\|x-y\|^3} \right) dy \\ = \int_{\Omega \setminus \Omega^\infty} \sigma^\infty \nabla \chi u^\infty \times \frac{x-y}{\|x-y\|^3} dy = \int_{\tilde{\Omega}} \sigma^\infty \nabla \chi u^\infty \times \frac{x-y}{\|x-y\|^3} dy. \end{aligned}$$

This shows that equation (2.50) is valid, which finishes the proof. \square

We have now derived FEM formulations for the MEG forward problem for both the direct approaches as well as for the local subtraction approach, where for the latter we assumed an isotropic source conductivity σ^∞ . Having derived the FEM approach, the next question is, of course, how to properly implement it. We will tackle this question next, starting with a discussion on how to assemble the local subtraction right-hand side.

2.6 Computing the local subtraction right-hand side

2.6.1 Introduction

In the local subtraction approach for a point electrode model [91], which contains the classical subtraction approach as described in [239] as a special case, one needs to compute integrals of

the form

$$I_P = \int_T \langle \sigma^c \nabla u^\infty, \nabla \varphi \rangle dV \quad (2.51)$$

$$I_S = \int_F \langle \sigma^\infty \nabla u^\infty, \eta \rangle \varphi dS \quad (2.52)$$

$$I_T = \int_T \langle \sigma \nabla \chi \cdot u^\infty, \nabla \varphi \rangle dV, \quad (2.53)$$

where we call I_P a *patch integral*, I_S a *surface integral*, and I_T a *transition integral*. When extending the local subtraction approach to also cover the complete electrode model, one additionally has to evaluate integrals of the form

$$I_I = \int_F \chi \cdot u^\infty \cdot \varphi dS \quad (2.54)$$

$$I_D = \int_F \chi \cdot u^\infty dS, \quad (2.55)$$

where we call I_I an *electrode interface integral*, and I_D an *electrode DOF integral*, see e.g. definition 2.5.12. Here, T is a mesh element and F is a facet of a mesh element.

To be able to derive a computationally viable FEM approach, we first need explicit formulas for u^∞ . Such expressions have already been known for a long time, both for isotropic σ^∞ and for anisotropic σ^∞ , see e.g. [239]. Note, however, that [239] cites [73] as a source for the anisotropic case. Now, while the author in [73] states a more general theorem, which contains the anisotropic case as a special case, he refers to his own diploma thesis for a proof of this general theorem. As far as we are aware, this diploma thesis is not readily available. To give a complete and self-contained discussion of the anisotropic case, and also because the formula for u^∞ is of fundamental importance for the local subtraction approach, we will thus include a derivation of u^∞ in this manuscript.

Once we have an expression for u^∞ , we then need to utilize it for the computation of the FEM right-hand side integrals (2.51), (2.52), (2.53), (2.54), and (2.55). Originally, this was done by either projecting u^∞ into a space of more easily integrable functions, see e.g. [239, 20], or by using numerical integration [48]. In 2019, Beltrachini published [18], where he was able to derive analytical expressions for the classical subtraction right-hand side integrals, i.e. (2.51) and (2.52), in the case of affine trial functions on tetrahedral meshes and isotropic σ^∞ . These formulas were later improved on, and in particular extended to the local subtraction case for point electrode models, by us in [91]. Here, we want to present these analytical expressions, as well as their extensions to the complete electrode model integrals (2.54) and (2.55). Additionally, we will extend the analytical expressions for the integrals (2.51)-(2.55) to the case of anisotropic conductivities. We want to emphasize that the analytical expressions are still restricted to tetrahedral meshes and piecewise affine trial functions.

This part of the thesis is structured as follows. In 2.6.2, we will first derive the infinity potential in both the isotropic and anisotropic cases. In 2.6.3, we will then derive the analytical expressions for the local subtraction right-hand side integrals under the assumption of an isotropic σ^∞ . In 2.6.4, we will then finally show how the analytical expressions for the local subtraction right-hand side integrals can be extended from the isotropic case to the anisotropic case.

For the remainder of this section, we will always assume that $T \subset \mathbb{R}^3$ is a tetrahedron and that $F \subset \mathbb{R}^3$ is a triangle. Furthermore, we will always assume that we employ piecewise affine test functions.

2.6.2 Deriving the infinity potential

The isotropic case

Our first goal is to solve the equation

$$\operatorname{div}(\sigma^\infty \nabla u^\infty) = \operatorname{div}(M \cdot \delta_{x_0}) \quad \text{on } \mathbb{R}^3$$

in the case of $\sigma^\infty = \operatorname{Id}$, where Id denotes the identity matrix. More concretely, we want to compute a solution of the equation

$$\Delta u^\infty = \operatorname{div}(M \cdot \delta_{x_0}) \quad \text{on } \mathbb{R}^3.$$

Using such a solution, we will then be able to deduce solutions for arbitrary symmetric positive definite $\sigma^\infty \in \mathbb{R}^{3 \times 3}$. The central tool in the construction of such a solution is the *Newton potential*

$$N(y, x) = -\frac{1}{4\pi} \frac{1}{\|x - y\|}.$$

It is well known, and has already been used multiple times in this thesis, that the function $N_y(x) = N(y, x)$ is a fundamental solution of the Laplace operator, i.e. $\Delta N_y = \delta_y$ on \mathbb{R}^3 . A proof of this can, e.g., be found in [213], 6.5(d).¹² Note that then for $1 \leq i \leq 3$ we have

$$\frac{\partial N}{\partial y_i}(y, x) = -\frac{1}{4\pi} \frac{x_i - y_i}{\|x - y\|^3}.$$

By standard theory, it is known that e.g. for $\rho \in C_c^2(\mathbb{R}^3)$ we have that

$$\Delta(N_0 * \rho) = \rho,$$

where $*$ denotes convolution, see e.g. [64], §16, Satz 2. We could now formally write out the definition of $N_0 * \operatorname{div}(M \cdot \delta_{x_0})$, and evaluate it using the standard “rules” for integrating the Dirac distribution. It turns out that this in fact produces the correct u^∞ . But $\operatorname{div}(M \cdot \delta_{x_0})$ is highly singular, and we are not aware of a version of the aforementioned theorem that has general enough prerequisites to be applicable here. We think that it is extremely likely that there is some publication where this case is covered. However, as it is relatively straightforward to extend the proof strategies from [64], §16, to rigorously derive the necessary properties of u^∞ , we decided to present our own short and self-contained proof in this thesis. In the following, we will use the notion of a *distribution*, which is sometimes also called a *generalized function*. There are many great resources for readers not familiar with this topic. We want to, in particular, recommend [170], Part II, for a comprehensive introduction with a heavy focus on functional analytic considerations, or [64], §17 (in German), for a more streamlined presentation that suffices for the present discussion.

Our first step consists of proving a technical lemma, which asserts that integrals against the Newton potential can be differentiated under the integral.

Lemma 2.6.1. *Let $\varphi \in C_c^\infty(\mathbb{R}^3)$. Define $g : \mathbb{R}^3 \rightarrow \mathbb{R}$ by*

$$g(y) = \int_{\mathbb{R}^3} N(y, x) \varphi(x) dx.$$

Then $g \in C^1(\mathbb{R}^3)$ and we have

$$\frac{\partial g}{\partial y_i}(y) = \int_{\mathbb{R}^3} \frac{\partial N}{\partial y_i}(y, x) \varphi(x) dx.$$

¹²Another excellent reference (in German) is [64], §17, Satz 2.

Proof. First note that $x \mapsto N(y, x)$ and $x \mapsto \frac{\partial N}{\partial y_i}(y, x)$ are locally integrable, and hence all integrals above are well-defined. Now we have

$$g(y) = \int_{\mathbb{R}^3} N(y, x) \varphi(x) dx \stackrel{z=x-y}{=} \int_{\mathbb{R}^3} -\frac{1}{4\pi} \frac{1}{\|z\|} \varphi(z+y) dz.$$

On a small environment of y , the integrand on the right-hand side is integrable as a function of z , differentiable as a function of y , and the derivatives have an L^1 -bound which does not depend on y . By [64], §11, Satz 12, g is hence differentiable in an environment of y . If now $R > 0$ is large enough, we have

$$\begin{aligned} \frac{\partial g}{\partial y_i}(y) &= \int_{\mathbb{R}^3} -\frac{1}{4\pi} \frac{1}{\|z\|} \frac{\partial}{\partial y_i} ((y, z) \mapsto \varphi(z+y)) dz \\ &= \int_{\mathbb{R}^3} -\frac{1}{4\pi} \frac{1}{\|z\|} \frac{\partial}{\partial z_i} ((y, z) \mapsto \varphi(z+y)) dz \\ &\stackrel{\varphi \in C_c^\infty(\mathbb{R}^3)}{=} \int_{B_R(0)} -\frac{1}{4\pi} \frac{1}{\|z\|} \frac{\partial}{\partial z_i} ((y, z) \mapsto \varphi(z+y)) dz \\ &= \lim_{\epsilon \rightarrow 0} \int_{B_R(0) \setminus B_\epsilon(0)} -\frac{1}{4\pi} \frac{1}{\|z\|} \frac{\partial}{\partial z_i} ((y, z) \mapsto \varphi(z+y)) dz \\ &\stackrel{\text{Gauss}}{=} \lim_{\epsilon \rightarrow 0} \int_{\partial B_\epsilon(0)} \frac{1}{4\pi} \frac{1}{\|z\|} \varphi(z+y) \eta_i dS(z) \\ &\quad + \int_{B_R(0) \setminus B_\epsilon(0)} \frac{1}{4\pi} \frac{\partial}{\partial z_i} \left(\frac{1}{\|z\|} \right) \varphi(z+y) dz, \end{aligned}$$

where η_i denotes the i -th component of the unit outer normal of $B_\epsilon(0)$. Since φ is bounded, we have for the first summand

$$\begin{aligned} \left| \int_{\partial B_\epsilon(0)} \frac{1}{4\pi} \frac{1}{\|z\|} \varphi(z+y) \eta_i dS(z) \right| &= \frac{1}{4\pi\epsilon} \left| \int_{\partial B_\epsilon(0)} \varphi(z+y) \eta_i dS(z) \right| \\ &\leq \frac{C}{4\pi\epsilon} \text{vol}(\partial B_\epsilon(0)) = C\epsilon \xrightarrow{\epsilon \rightarrow 0} 0. \end{aligned}$$

Since $\frac{\partial}{\partial z_i} \left(\frac{1}{\|z\|} \right) = -\frac{z_i}{\|z\|^3}$, this implies

$$\frac{\partial g}{\partial y_i}(y) = \int_{\mathbb{R}^3} -\frac{1}{4\pi} \frac{z_i}{\|z\|^3} \varphi(z+y) dz \stackrel{z=x-y}{=} \int_{\mathbb{R}^3} -\frac{1}{4\pi} \frac{x_i - y_i}{\|x - y\|^3} \varphi(x) dx = \int_{\mathbb{R}^3} \frac{\partial N}{\partial y_i}(y, x) \varphi(x) dx.$$

This finishes the proof. \square

Now define for $x_0, M \in \mathbb{R}^3$

$$f_{x_0, M}(x) := -\langle M, \nabla_y N(x_0, x) \rangle,$$

where $\nabla_y N(y, x) = \left(\frac{\partial N}{\partial y_1}(y, x), \frac{\partial N}{\partial y_2}(y, x), \frac{\partial N}{\partial y_3}(y, x) \right)$. We will now show that $f_{x_0, M}$ solves the dipole problem for $\sigma^\infty = 1$.

Theorem 2.6.2. *Let $f_{x_0, M}$ be as above. Then*

$$\Delta f_{x_0, M} = \text{div}(M \cdot \delta_{x_0}).$$

Proof. Let $\varphi \in C_c^\infty(\mathbb{R}^3)$. Then

$$\operatorname{div}(M \cdot \delta_{x_0})(\varphi) = \sum_{i=1}^3 \frac{\partial}{\partial x_i} (M_i \delta_{x_0})(\varphi) = \sum_{i=1}^3 -M_i \delta_{x_0} \left(\frac{\partial \varphi}{\partial x_i} \right) = -\langle M, \nabla \varphi(x_0) \rangle.$$

Furthermore, we can compute

$$\begin{aligned} \Delta f_{x_0, M}(\varphi) &= \int_{\mathbb{R}^3} f_{x_0, M}(x) \Delta \varphi(x) dx = - \sum_{i=1}^3 M_i \int_{\mathbb{R}^3} \frac{\partial N}{\partial y_i}(x_0, x) \Delta \varphi(x) dx \\ &\stackrel{\text{Lemma 2.6.1}}{=} - \sum_{i=1}^3 M_i \frac{\partial}{\partial y_i} \left(y \mapsto \int_{\mathbb{R}^3} N(y, x) \Delta \varphi(x) dx \right) (x_0) \\ &\stackrel{\Delta N_{y=\delta_y}}{=} - \sum_{i=1}^3 M_i \frac{\partial \varphi}{\partial y_i}(x_0) = -\langle M, \nabla \varphi(x_0) \rangle = \operatorname{div}(M \cdot \delta_{x_0})(\varphi). \end{aligned}$$

□

We have thus shown that $f_{x_0, M}$ is a solution of the dipole problem for $\sigma^\infty = \operatorname{Id}$. If $\sigma^\infty = \lambda \cdot \operatorname{Id}$ with $\lambda > 0$, we immediately see that a solution of $\operatorname{div}(\sigma^\infty \nabla u^\infty) = \operatorname{div}(M \cdot \delta_{x_0})$ is given by $\frac{1}{\lambda} f_{x_0, M}$. If $\sigma^\infty = \lambda \cdot \operatorname{Id}$, we thus see that

$$u^\infty(x) = \frac{1}{4\pi\lambda} \frac{\langle M, x - x_0 \rangle}{\|x - x_0\|^3} \quad (2.56)$$

is a solution of $\operatorname{div}(\sigma^\infty \nabla u^\infty) = \operatorname{div}(M \cdot \delta_{x_0})$ on \mathbb{R}^3 .

Note, however, that σ^∞ can be an arbitrary positive definite matrix. In the next subsection, we will show how a solution of the anisotropic problem can be derived from a solution of the isotropic problem.

The anisotropic case

We now assume that $\sigma^\infty \in \mathbb{R}^{3 \times 3}$ is an arbitrary symmetric positive definite matrix. We will construct a solution of the corresponding dipole problem via a suitable change of coordinates. First, note that since σ^∞ is symmetric positive definite, its inverse $(\sigma^\infty)^{-1}$ is also symmetric positive definite, and it can thus be decomposed as

$$(\sigma^\infty)^{-1} = A \cdot A^\top,$$

where $A \in \mathbb{R}^{3 \times 3}$ and $\det(A) > 0$. The matrix A can e.g. be chosen as a matrix square root of $(\sigma^\infty)^{-1}$, or alternatively from a Cholesky decomposition of $(\sigma^\infty)^{-1}$.

Now, define $\Phi_{A^\top} : \mathbb{R}^3 \rightarrow \mathbb{R}^3; x \mapsto A^\top \cdot x$. We will now show that the transformation Φ_{A^\top} reduces the anisotropic dipole problem to an isotropic dipole problem. The following argument is inspired by [64], §10, Satz 4, where the author investigates how the Laplace operator transforms under arbitrary coordinate changes.

In the following, if M is an arbitrary matrix, we define $\Phi_M(x) := M \cdot x$. We start with a simple application of the chain rule, which we will make extensive use of.

Lemma 2.6.3. *Let $U \subset \mathbb{R}^3$ be open. Let $M \in \mathbb{R}^{3 \times 3}$ be an arbitrary matrix, and let $g : \Phi_M(U) \rightarrow \mathbb{R}$ be differentiable. For $y \in U$ we then have*

$$\nabla (g \circ \Phi_M)(y) = M^\top \cdot (\nabla g) \circ \Phi_M(y). \quad (2.57)$$

$$(\nabla g) \circ \Phi_M(y) = M^{-\top} \cdot \nabla (g \circ \Phi_M)(y). \quad (2.58)$$

Proof. Note that $\nabla g = J_g^\top$, where J_g denotes the jacobian of g . By the chain rule, we thus have

$$\nabla (g \circ \Phi_M)(y) = J_{g \circ \Phi_M}^\top(y) = (J_g(\Phi_M(y)) \cdot J_{\Phi_M}(y))^\top = M^\top \cdot \nabla g \circ \Phi_M(y).$$

This shows equation 2.57. Multiplying by $M^{-\top}$ shows equation 2.58. \square

We will now quickly repeat a special case of the transformation formula, as we will use it numerous times in the following.

Theorem 2.6.4 (Transformation formula, linear case). *Let $g : \mathbb{R}^3 \rightarrow \mathbb{R}$ be integrable and let $M \in \mathbb{R}^{3 \times 3}$ be invertible. Then $g \circ \Phi_M$ is also integrable over \mathbb{R}^3 , and we have*

$$\int_{\mathbb{R}^3} g dV = \int_{\mathbb{R}^3} (g \circ \Phi_M) \cdot |\det(M)| dV$$

Proof. Noting that $\mathbb{R}^3 = \Phi_M(\mathbb{R}^3)$ and $J_{\Phi_M} = M$, it follows that this theorem is just a special case of the general transformation formula, which can e.g. be found in [64], §9, Satz 2. \square

Lemma 2.6.5. *Let $\sigma^\infty \in \mathbb{R}^{3 \times 3}$ be symmetric positive definite, and let $A \in \mathbb{R}^{3 \times 3}$ be such that $(\sigma^\infty)^{-1} = A \cdot A^\top$ and $\det(A) > 0$. We then have for an arbitrary $\varphi \in C^2(\mathbb{R}^3)$ that*

$$\operatorname{div}(\sigma^\infty \nabla \varphi) \circ \Phi_{A^{-\top}} = \Delta(\varphi \circ \Phi_{A^{-\top}}). \quad (2.59)$$

Proof. By the fundamental lemma of the calculus of variations, it suffices to show that for all test functions $\xi \in C_c^\infty(\mathbb{R}^3)$ we have

$$\int_{\mathbb{R}^3} \operatorname{div}(\sigma^\infty \nabla \varphi) \circ \Phi_{A^{-\top}} \cdot \xi dV = \int_{\mathbb{R}^3} \Delta(\varphi \circ \Phi_{A^{-\top}}) \cdot \xi dV.$$

Note that, since ξ has compact support, boundary terms vanish when applying partial integration on a sufficiently large ball. We can now compute

$$\begin{aligned} \int_{\mathbb{R}^3} \operatorname{div}(\sigma^\infty \nabla \varphi) \circ \Phi_{A^{-\top}} \cdot \xi dV &= \int_{\mathbb{R}^3} \operatorname{div}(\sigma^\infty \nabla \varphi) \cdot \xi \circ \Phi_{A^\top} \cdot \det(A^\top) dV \\ &= - \int_{\mathbb{R}^3} \langle \sigma^\infty \nabla \varphi, \nabla(\xi \circ \Phi_{A^\top}) \rangle \cdot \det(A^\top) dV \\ &= - \int_{\mathbb{R}^3} \langle \sigma^\infty \nabla \varphi, A(\nabla \xi) \circ \Phi_{A^\top} \rangle \cdot \det(A^\top) dV \\ &= - \int_{\mathbb{R}^3} \langle A^\top \sigma^\infty \nabla \varphi, (\nabla \xi) \circ \Phi_{A^\top} \rangle \cdot \det(A^\top) dV \\ &= - \int_{\mathbb{R}^3} \langle A^\top \sigma^\infty (\nabla \varphi) \circ \Phi_{A^{-\top}}, \nabla \xi \rangle dV \\ &= \int_{\mathbb{R}^3} \operatorname{div}(A^\top \sigma^\infty (\nabla \varphi) \circ \Phi_{A^{-\top}}) \cdot \xi dV. \end{aligned}$$

Now, note that $\sigma^\infty = A^{-\top} \cdot A^{-1}$, and thus

$$A^\top \sigma^\infty (\nabla \varphi) \circ \Phi_{A^{-\top}} = A^{-1} (\nabla \varphi) \circ \Phi_{A^{-\top}} = \nabla (\varphi \circ \Phi_{A^{-\top}}),$$

which finishes the proof. \square

We can now finally derive the anisotropic infinity potential.

Theorem 2.6.6. *Let $\sigma^\infty \in \mathbb{R}^{3 \times 3}$ be symmetric positive definite, and let $A \in \mathbb{R}^{3 \times 3}$ be such that $(\sigma^\infty)^{-1} = A \cdot A^\top$ and $\det(A) > 0$. Let $x_0 \in \mathbb{R}^3$ be a dipole position and $M \in \mathbb{R}^3$ be a dipole moment. Denote by*

$$f_{A^\top x_0, A^\top M}(x) := \frac{1}{4\pi} \frac{\langle A^\top M, x - A^\top x_0 \rangle}{\|x - A^\top x_0\|^3}$$

a solution of $\Delta f_{A^\top x_0, A^\top M} = \operatorname{div}(A^\top M \cdot \delta_{A^\top x_0})$, as shown in theorem 2.6.2. Then, the function

$$u_{x_0, M, \sigma^\infty}^\infty := \det(A) \cdot f_{A^\top x_0, A^\top M} \circ \Phi_{A^\top}$$

solves $\operatorname{div}(\sigma^\infty \nabla u_{x_0, M, \sigma^\infty}^\infty) = \operatorname{div}(M \cdot \delta_{x_0})$ in the distributional sense, i.e. $u_{x_0, M, \sigma^\infty}^\infty$ is a solution of the dipole problem for anisotropic σ^∞ . Furthermore, $u_{x_0, M, \sigma^\infty}^\infty$ can be evaluated without needing to compute A via

$$u_{x_0, M, \sigma^\infty}^\infty(x) = \frac{1}{4\pi \sqrt{\det(\sigma^\infty)}} \frac{\langle M, (\sigma^\infty)^{-1}(x - x_0) \rangle}{\langle (\sigma^\infty)^{-1}(x - x_0), x - x_0 \rangle^{\frac{3}{2}}}$$

Proof. First, note that $u_{x_0, M, \sigma^\infty}^\infty$ is locally integrable, and hence defines a distribution. Now let $\varphi \in C_c^\infty(\mathbb{R}^3)$ be an arbitrary test function. We then have by lemma 2.6.5 that

$$\begin{aligned} \operatorname{div}(\sigma^\infty \nabla u_{x_0, M, \sigma^\infty}^\infty)(\varphi) &= \int_{\mathbb{R}^3} (\det(A) \cdot f_{A^\top x_0, A^\top M} \circ \Phi_{A^\top}) \cdot \operatorname{div}(\sigma^\infty \nabla \varphi) \, dV \\ &= \int_{\mathbb{R}^3} f_{A^\top x_0, A^\top M} \cdot \operatorname{div}(\sigma^\infty \nabla \varphi) \circ \Phi_{A^{-\top}} \, dV \\ &= \int_{\mathbb{R}^3} f_{A^\top x_0, A^\top M} \cdot \Delta(\varphi \circ \Phi_{A^{-\top}}) \, dV \\ &\stackrel{\text{Theorem 2.6.2}}{=} -\langle A^\top M, \nabla(\varphi \circ \Phi_{A^{-\top}})(A^\top x_0) \rangle. \end{aligned}$$

Now, we have

$$\langle A^\top M, \nabla(\varphi \circ \Phi_{A^{-\top}})(A^\top x_0) \rangle = \langle M, A \cdot \nabla(\varphi \circ \Phi_{A^{-\top}})(A^\top x_0) \rangle = \langle M, \nabla \varphi(x_0) \rangle,$$

and thus

$$\operatorname{div}(\sigma^\infty \nabla u_{x_0, M, \sigma^\infty}^\infty)(\varphi) = \operatorname{div}(M \cdot \delta_{x_0})(\varphi).$$

Thus, $u_{x_0, M, \sigma^\infty}^\infty$ is a solution of the anisotropic dipole problem on \mathbb{R}^3 . Noting that

$$\det(\sigma^\infty) = \frac{1}{\det((\sigma^\infty)^{-1})} = \frac{1}{\det(A)^2} \iff \det(A) = \frac{1}{\sqrt{\det(\sigma^\infty)}},$$

we can now finally compute

$$\begin{aligned} \det(A) \cdot f_{A^\top x_0, A^\top M} \circ \Phi_{A^\top}(x) &= \frac{1}{4\pi \sqrt{\det(\sigma^\infty)}} \cdot \frac{\langle A^\top M, A^\top(x) - A^\top x_0 \rangle}{\|A^\top x - A^\top x_0\|^3} \\ &= \frac{1}{4\pi \sqrt{\det(\sigma^\infty)}} \cdot \frac{\langle M, (\sigma^\infty)^{-1}(x - x_0) \rangle}{\langle (\sigma^\infty)^{-1}(x - x_0), x - x_0 \rangle^{\frac{3}{2}}}. \end{aligned}$$

This finishes the proof. \square

2.6.3 Right-hand side analytical expressions in the isotropic case

Here, we aim to give a self-contained derivation of the analytical expressions for the integrals 2.51, 2.52, 2.53, 2.54, and 2.55. We want to emphasize that the derivation of the integrals 2.51 and 2.52 is due to Beltrachini [18], based on prior work of Graglia [75] and Wilton [235]. Our contribution consists of the derivation of analytical expressions for the integrals 2.53, 2.54, and 2.55. Similar to [18], we also make heavy use of [75] and [235].

We have decided to give a self-contained derivation for three reasons. First, the tools needed for our new derivations of 2.53, 2.54, and 2.55 almost suffice to also derive 2.51 and 2.52, i.e. including the derivations of 2.51 and 2.52 does not render the argument substantially more complicated. Secondly, all formulas derived below are implemented into the DUNEuro toolbox. We think that the understandability and maintainability of the code greatly increase when there is a single reference for its derivation that also matches the implementation, instead of forcing a prospective maintainer to piece together the necessary information from at least four different papers with mutually different notations and conventions. Thirdly, we have substantially improved on the derivation. On the one hand, we show how numerous derivations in [235], [75], and [18] can be reduced to a handful of core computations, which can also be applied to a larger class of integral problems. On the other hand, we improved the numerical stability of the analytical expressions, as, for the previous formulas, there were edge cases where certain combinations of the position of the dipole and the position of the mesh element led to singularities in the formulas that could completely destroy the simulation.¹³

Describing the geometry of a triangle

Let $F \subset \mathbb{R}^3$ denote a triangle with corners p_1, p_2 , and p_3 . As a first step, we want to explicitly describe a map of F . The results of this subsection are due to [235] and [75]. The proof of lemma 2.6.7 is due to us. Additionally, as [235] and [75] mainly only state the final formulas, the strategies presented below to perform the computations are also due to us. If the reader is interested in following along with the derivation, we highly recommend drawing a figure.

First, we define

$$\begin{aligned} s_1 &= \frac{p_3 - p_2}{\|p_3 - p_2\|}, \gamma_1 = \|p_3 - p_2\| \\ s_2 &= \frac{p_1 - p_3}{\|p_1 - p_3\|}, \gamma_2 = \|p_1 - p_3\| \\ s_3 &= \frac{p_2 - p_1}{\|p_2 - p_1\|}, \gamma_3 = \|p_2 - p_1\| \end{aligned}$$

Now, define $u := s_3$, $w = \frac{s_1 \times s_2}{\|s_1 \times s_2\|}$, and $v = w \times u$, where \times denotes the cross product. Note that w is a unit normal vector to the plane defined by the triangle F . Furthermore, for $1 \leq i \leq 3$, let $m_i \in \mathbb{R}^3$ be the unique normal vector to the edge given by s_i of the triangle F in the plane defined by F that is pointing outside of the triangle.

Lemma 2.6.7. *We have the following.*

1. (u, v, w) is a right-handed orthonormal basis of \mathbb{R}^3 .

¹³In fact, after having spent two months studying [235], [75], and [18], and implementing the analytical expressions for the classical subtraction approach into DUNEuro, this happened for the very first dipole I tested. Imagine my dejection.

$$2. v_3 := \langle v, p_3 - p_1 \rangle > 0$$

3. For $1 \leq i \leq 3$ we have $m_i = s_i \times w$.

Proof. If $a, b, c \in \mathbb{R}^3$, then we have $\det(a, b, c) = \langle a, b \times c \rangle$. Furthermore, if (a, b, c) is a basis of \mathbb{R}^3 , we say that the basis is right-handed if $\det(a, b, c) > 0$. In particular, if a and b are linearly independent we always have

$$\det(a, b, a \times b) = \det(a \times b, a, b) = \langle a \times b, a \times b \rangle > 0,$$

i.e. $(a, b, a \times b)$ is right-handed. Furthermore, note that we have

$$\gamma_3 s_3 = p_2 - p_1 = (p_2 - p_3) + (p_3 - p_1) = -\gamma_1 s_1 - \gamma_2 s_2,$$

and $\gamma_1, \gamma_2, \gamma_3 > 0$. Since $a \times b = -b \times a$, we can now compute

$$\det(u, v, w) = \det(u, w \times u, w) = \det(u, w, u \times w) > 0,$$

i.e. (u, v, w) is right-handed. Furthermore, since s_3 is in the plane defined by s_1 and s_2 , w is orthogonal to u , and furthermore, v is also orthogonal to u and w . Finally, since for $a, b \in \mathbb{R}^3$ we have

$$\|a \times b\|^2 = \|a\|^2 \|b\|^2 - \langle a, b \rangle^2,$$

it follows that $\|u\| = \|w\| = \|v\| = 1$. Thus, (u, v, w) is a right-handed orthonormal basis of \mathbb{R}^3 . Furthermore, we have

$$\begin{aligned} \text{sign}(\langle v, p_3 - p_1 \rangle) &= -\text{sign}(\langle v, s_2 \rangle) = -\text{sign}(\langle (s_1 \times s_2) \times s_3, s_2 \rangle) \\ &= -\text{sign}(\det(s_2, s_1 \times s_2, s_3)) \\ &= -\text{sign}(\det(s_2, s_1 \times s_2, -\gamma_1 s_1 - \gamma_2 s_2)) \\ &= \text{sign}(\det(s_2, s_1 \times s_2, s_1)) \\ &= \text{sign}(\det(s_1, s_2, s_1 \times s_2)) = 1, \end{aligned}$$

and hence $\langle v, p_3 - p_1 \rangle > 0$.

Finally, note that for $1 \leq i \leq 3$ we have that m_i is by definition orthogonal to s_i , and, since it is contained in the plane defined by F , we have that m_i is also orthogonal to w . Since s_i and w are orthogonal, we thus have $m_i = \pm s_i \times w$.

Since m_1 by definition points in the opposite direction of $p_1 - p_3$, we have that $\langle m_1, p_1 - p_3 \rangle < 0$. Since

$$\text{sign}(\langle s_1 \times w, p_1 - p_3 \rangle) = \text{sign}(\langle s_1 \times w, s_2 \rangle) = \text{sign}(\langle s_1 \times (s_1 \times s_2), s_2 \rangle) = -1,$$

and hence $\langle s_1 \times w, p_1 - p_3 \rangle < 0$ and thus $m_1 = s_1 \times w$. The equations $m_2 = s_2 \times w$ and $m_3 = s_3 \times w$ follow similarly. \square

Now, define $O = (u, v, w) \in \mathbb{R}^{3 \times 3}$, and $\begin{pmatrix} u_3 \\ v_3 \\ w_3 \end{pmatrix} = O^\top \cdot (p_3 - p_1)$. Note that $p_3 - p_1$ is in the plane defined by the triangle F , and hence $w_3 = 0$. Furthermore, by lemma 2.6.7 we have $v_3 > 0$. If we now define $q_1 = (0, 0)$, $q_2 = (\gamma_3, 0)$, and $q_3 = (u_3, v_3) \in \mathbb{R}^2$, we have that

$\widehat{F}_0 = \text{conv}\{q_1, q_2, q_3\} \subset \mathbb{R}^2$ forms a triangle. Here, conv denotes the convex hull. If we now define

$$\mu_F : \widehat{F}_0 \rightarrow F; \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \mapsto (u, v) \cdot \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} + p_1, \quad (2.60)$$

we get an affine map of the triangle F with $\mu_F(q_i) = p_i$ for $1 \leq i \leq 3$. Furthermore, note that since u and v are orthonormal, the gramian of this map is 1 and we thus have for $g : F \rightarrow \mathbb{R}$ integrable that

$$\int_F g \, dS = \int_{\widehat{F}_0} g \circ \mu_F \, d\lambda_2,$$

where λ_2 denotes the 2-dimensional Lebesgue measure. Additionally, note that the inverse of the map is given by

$$\mu_F^{-1} : \text{aff}(F) \rightarrow \mathbb{R}^2; x \mapsto \begin{pmatrix} \langle u, x - p_1 \rangle \\ \langle v, x - p_1 \rangle \end{pmatrix}. \quad (2.61)$$

In the following, we will always assume that we have a dipole at position $x_0 \in \mathbb{R}^3$ with moment $M \in \mathbb{R}^3$. Furthermore, we will assume that $x_0 \notin F$. Let $\rho \in \mathbb{R}^3$ be the orthogonal projection of x_0 onto $\text{aff}(F)$. Define

$$\begin{pmatrix} u_0 \\ v_0 \end{pmatrix} := \mu_F^{-1}(\rho),$$

and $w_0 := -\langle w, x_0 - p_1 \rangle$. Then, $\rho = x_0 + w_0 \cdot w = p_1 + u_0 \cdot u + v_0 \cdot v$.

Now, let $\widehat{m}_i \in \mathbb{R}^2$ denote the outer normal of the edge of the triangle \widehat{F}_0 corresponding to the edge given by s_i of the triangle F . Since, by lemma 2.6.7, traversing the triangle from q_1 to q_2 to q_3 and back to q_1 corresponds to a counterclockwise orientation of the edges of \widehat{F}_0 , we can obtain the outer normal direction by a 90° clockwise rotation of the differences $q_2 - q_1$, $q_3 - q_2$, and $q_1 - q_3$. Since the edge lengths of the edges of \widehat{F}_0 are the same as the edge lengths of F , we thus get

$$\widehat{m}_1 = \frac{1}{\gamma_1} \begin{pmatrix} v_3 \\ \gamma_3 - u_3 \end{pmatrix}, \widehat{m}_2 = \frac{1}{\gamma_2} \begin{pmatrix} -v_3 \\ u_3 \end{pmatrix}, \widehat{m}_3 = \begin{pmatrix} 0 \\ -1 \end{pmatrix}.$$

Furthermore, note that $(u, v) \cdot \widehat{m}_i = m_i$.

Now, for $1 \leq i \leq 3$, let $t_i \in \mathbb{R}$ denote the unique scalar such that the vector $\rho + t_i \cdot m_i$ lies on the line defined by the edge corresponding to s_i . Let $p_4 := p_1$ and $q_4 := q_1$. By elementary geometry, the orthogonal projection of ρ onto this line is given by

$$p_{i+1} + \langle \rho - p_{i+1}, s_i \rangle \cdot s_i,$$

and hence we have

$$p_{i+1} + \langle \rho - p_{i+1}, s_i \rangle \cdot s_i = \rho + t_i \cdot m_i,$$

which, after multiplication with m_i , implies that

$$t_i = \langle p_{i+1} - \rho, m_i \rangle = \langle (u, v)^\top (p_{i+1} - \rho), \widehat{m}_i \rangle = \langle q_{i+1} - \begin{pmatrix} u_0 \\ v_0 \end{pmatrix}, \widehat{m}_i \rangle.$$

A direct computation then shows that

$$\begin{aligned} t_1 &= \frac{v_0(u_3 - \gamma_3) + v_3(\gamma_3 - u_0)}{\gamma_1} \\ t_2 &= \frac{u_0 v_3 - v_0 u_3}{\gamma_2} \\ t_3 &= v_0. \end{aligned}$$

Now, let $p_5 := p_2$ and $q_5 := q_2$. Then, for $1 \leq i \leq 3$ let γ_i^- denote the unique scalar such that $p_{i+1} = \rho + t_i \cdot m_i + \gamma_i^- \cdot s_i$, and let γ_i^+ denote the unique scalar such that $p_{i+2} = \rho + t_i \cdot m_i + \gamma_i^+ \cdot s_i$. Similar to the argument above, multiplication by s_i shows that we have

$$\gamma_i^- = \langle p_{i+1} - \rho, s_i \rangle = \frac{1}{\gamma_i} \cdot \langle q_{i+1} - \begin{pmatrix} u_0 \\ v_0 \end{pmatrix}, q_{i+2} - q_{i+1} \rangle$$

and

$$\gamma_i^+ = \frac{1}{\gamma_i} \cdot \langle q_{i+2} - \begin{pmatrix} u_0 \\ v_0 \end{pmatrix}, q_{i+2} - q_{i+1} \rangle.$$

Directly computing these expressions yields

$$\begin{aligned} \gamma_1^- &= -\frac{(\gamma_3 - u_0)(\gamma_3 - u_3) + v_0 v_3}{\gamma_1} \\ \gamma_1^+ &= \frac{(u_3 - u_0)(u_3 - \gamma_3) + v_3(v_3 - v_0)}{\gamma_1} \\ \gamma_2^- &= -\frac{u_3(u_3 - u_0) + v_3(v_3 - v_0)}{\gamma_2} \\ \gamma_2^+ &= \frac{u_0 u_3 + v_0 v_3}{\gamma_2} \\ \gamma_3^- &= -u_0 \\ \gamma_3^+ &= \gamma_3 - u_0. \end{aligned}$$

To avoid potential confusion, we want to emphasize that the author of [18] makes a small mistake in his expression for γ_2^- .

If now for $a, b \in \mathbb{R}^k$ we denote by $[a, b]$ the line segment from a to b , we have for $1 \leq i \leq 3$ and an integrable function $g : [p_{i+1}, p_{i+2}] \rightarrow \mathbb{R}$ that

$$\int_{[p_{i+1}, p_{i+2}]} g \, ds = \int_{[q_{i+1}, q_{i+2}]} g \circ \mu_F \, ds \quad (2.62)$$

$$= \int_{\gamma_i^-}^{\gamma_i^+} g(\rho + t_i \cdot m_i + t \cdot s_i) \, dt. \quad (2.63)$$

Finally, let R_i denote the distance of x_0 from the line defined by the segment $[p_{i+1}, p_{i+2}]$, R_i^- the distance of x_0 from p_{i+1} and R_i^+ the distance of x_0 from p_{i+2} . By the Pythagorean theorem, we then have

$$\begin{aligned} R_i &= \sqrt{w_0^2 + t_i^2} \\ R_i^\pm &= \sqrt{w_0^2 + t_i^2 + (\gamma_i^\pm)^2}. \end{aligned}$$

Now, denote for $1 \leq i \leq 3$ by $\varphi_i : F \rightarrow \mathbb{R}$ the unique affine map such that $\varphi_i(p_i) = 1$, and $\varphi = 0$ on the remaining two vertices of the triangle. Then the composition $\varphi_i \circ \mu_F$ is also affine, and hence uniquely determined by its values on the set $\{q_1, q_2, q_3\}$. A direct computation then shows that for $(z_1, z_2) \in \mathbb{R}^2$ we have

$$\begin{pmatrix} \varphi_1 \\ \varphi_2 \\ \varphi_3 \end{pmatrix} \circ \mu_F(z_1, z_2) = \begin{pmatrix} 1 & -\frac{1}{\gamma_3} & \frac{\frac{u_3}{\gamma_3} - 1}{v_3} \\ 0 & \frac{1}{\gamma_3} & -\frac{u_3}{\gamma_3 v_3} \\ 0 & 0 & \frac{1}{v_3} \end{pmatrix} \cdot \begin{pmatrix} 1 \\ z_1 \\ z_2 \end{pmatrix} \quad (2.64)$$

We now have all the necessary prerequisites for computing integrals over F .

Computing integrals over a triangle in space

In this section, we use the same notation as in the “Describing the geometry of a triangle” section. Furthermore, we define

$$R : \mathbb{R}^3 \setminus \{x_0\} \rightarrow \mathbb{R}; x \mapsto \|x - x_0\|.$$

Our first goal will be to compute $\int_F R^m dS$ for certain values of $m \in \mathbb{Z}$. For $m = -1$, this value was computed by [235], for $m = -3$ by [75], and for $m = -5$ by [18]. Here, we will present a unified treatment for these integrals, which encompasses the previous derivations.

Lemma 2.6.8. *Let*

$$h : \mathbb{R}^2 \setminus \{(u_0, v_0)\} \rightarrow \mathbb{R}^2; (z_1, z_2) \mapsto \frac{1}{2} \cdot \frac{1}{(z_1 - u_0)^2 + (z_2 - v_0)^2} \cdot \begin{pmatrix} z_1 - u_0 \\ z_2 - v_0 \end{pmatrix}.$$

Furthermore, let $g : \mathbb{R}^2 \setminus \{(u_0, v_0)\} \rightarrow \mathbb{R}$ have a factorization $g = g_0 \circ d$, where $d(z_1, z_2) = (z_1 - u_0)^2 + (z_2 - v_0)^2$ and $g_0 : \mathbb{R} \rightarrow \mathbb{R}$ is an arbitrary function. If then $G_0 : \mathbb{R} \rightarrow \mathbb{R}$ is such that $G'_0 = g_0$, we have

$$g = \operatorname{div}((G_0 \circ d) \cdot h).$$

Proof. Astute readers have already noticed that h is a scalar multiple of the gradient of the 2-dimensional Newton potential, and hence $\operatorname{div}(h) = 0$ on $\mathbb{R}^2 \setminus \{(u_0, v_0)\}$. Alternatively, this statement can be verified by direct computation. We thus have

$$\operatorname{div}((G_0 \circ d) \cdot h) = \langle \nabla(G_0 \circ d), h \rangle.$$

Now,

$$\nabla(G_0 \circ d)(z_1, z_2) = J_d(z_1, z_2)^\top \cdot G'_0(d(z_1, z_2)) = 2 \cdot g(z_1, z_2) \cdot \begin{pmatrix} z_1 - u_0 \\ z_2 - v_0 \end{pmatrix},$$

and hence $\langle \nabla(G_0 \circ d), h \rangle = g$. □

Lemma 2.6.9. *Let $m \in \mathbb{Z}$ with $m \neq -2$ and let $w_0 \neq 0$, $t_1 \neq 0$, $t_2 \neq 0$, and $t_3 \neq 0$. Then*

$$\int_F R^m dS = \frac{1}{m+2} \cdot \sum_{i=1}^3 \int_{\gamma_i^-}^{\gamma_i^+} \left(\sqrt{w_0^2 + t_i^2 + t^2}^{m+2} - \sqrt{w_0^2}^{m+2} \right) \cdot \frac{t_i}{t_i^2 + t^2} dt.$$

Proof. First, note that

$$\begin{aligned} \int_F R^m dS &= \int_{\widehat{F}_0} R^m \circ \Phi(z_1, z_2) d\lambda_2(z_1, z_2) \\ &= \int_{\widehat{F}_0} \sqrt{w_0^2 + (z_1 - u_0)^2 + (z_2 - v_0)^2}^m d\lambda_2(z_1, z_2). \end{aligned}$$

If we now define $g_0(a) = \sqrt{w_0^2 + a}^m$, then the integrand in the last integral above is of the form $g_0 \circ d$ as defined in lemma 2.6.8. If we then define $G_0(a) := \frac{2}{m+2} \sqrt{w_0^2 + a}^{m+2}$, then $G'_0 = g_0$, and we have

$$\int_F R^m dS = \int_{\widehat{F}_0} \operatorname{div} \left(\frac{1}{2} \frac{G_0 \circ d(z_1, z_2)}{(z_1 - u_0)^2 + (z_2 - v_0)^2} \cdot \begin{pmatrix} z_1 - u_0 \\ z_2 - v_0 \end{pmatrix} \right) d\lambda_2(z_1, z_2)$$

We differentiate between two cases. First, if $(u_0, v_0) \notin \widehat{F}_0$. Then, the expression inside the divergence in the last integral is a well-defined smooth function in \widehat{F}_0 , and we have by the divergence theorem that

$$\begin{aligned} \int_{\widehat{F}_0} \operatorname{div} \left(\frac{1}{2} \frac{G_0 \circ d(z_1, z_2)}{(z_1 - u_0)^2 + (z_2 - v_0)^2} \cdot \begin{pmatrix} z_1 - u_0 \\ z_2 - v_0 \end{pmatrix} \right) d\lambda_2(z_1, z_2) \\ = \sum_{i=1}^3 \int_{[q_{i+1}, q_{i+2}]} \frac{1}{2} \frac{G_0 \circ d(z_1(t), z_2(t))}{(z_1(t) - u_0)^2 + (z_2(t) - v_0)^2} \cdot \left\langle \begin{pmatrix} z_1(t) - u_0 \\ z_2(t) - v_0 \end{pmatrix}, \widehat{m}_i \right\rangle ds(t) \\ = \frac{1}{m+2} \cdot \sum_{i=1}^3 \int_{[p_{i+1}, p_{i+2}]} \frac{R^{m+2}}{\|x - \rho\|^2} \langle x - \rho, m_i \rangle ds, \end{aligned}$$

where x denotes the identity function on the edge. By equation 2.63 we now have for $1 \leq i \leq 3$ that

$$\int_{[p_{i+1}, p_{i+2}]} \frac{R^{m+2}}{\|x - \rho\|^2} \langle x - \rho, m_i \rangle ds = \int_{\gamma_i^-}^{\gamma_i^+} \frac{\sqrt{w_0^2 + t_i^2 + t^2}^{m+2}}{t_i^2 + t^2} \cdot t_i dt$$

Furthermore, we have in this case

$$\begin{aligned} \sum_{i=1}^3 \int_{\gamma_i^-}^{\gamma_i^+} \frac{t_i}{t_i^2 + t^2} dt &= \int_{\partial \widehat{F}_0} \left\langle \frac{1}{(z_1(t) - u_0)^2 + (z_2(t) - v_0)^2} \begin{pmatrix} z_1(t) - u_0 \\ z_2(t) - v_0 \end{pmatrix}, \widehat{m} \right\rangle ds(t) \\ &= \int_{\widehat{F}_0} \operatorname{div} \left(\frac{1}{(z_1 - u_0)^2 + (z_2 - v_0)^2} \begin{pmatrix} z_1 - u_0 \\ z_2 - v_0 \end{pmatrix} \right) d\lambda_2(z_1, z_2) \\ &= 0, \end{aligned}$$

where we have again used that $\operatorname{div} \left(\frac{1}{(z_1 - u_0)^2 + (z_2 - v_0)^2} \begin{pmatrix} z_1 - u_0 \\ z_2 - v_0 \end{pmatrix} \right) = 0$. Putting this together, we have

$$\int_F R^m dS = \frac{1}{m+2} \cdot \sum_{i=1}^3 \int_{\gamma_i^-}^{\gamma_i^+} \left(\sqrt{w_0^2 + t_i^2 + t^2}^{m+2} - \sqrt{w_0^2}^{m+2} \right) \cdot \frac{t_i}{t_i^2 + t^2} dt.$$

Now we consider the case where (u_0, v_0) is in the interior of \widehat{F}_0 . Note that the conditions $t_1 \neq 0$, $t_2 \neq 0$, and $t_3 \neq 0$ imply that (u_0, v_0) cannot lie on the boundary of \widehat{F}_0 . Now, denote by $B_\epsilon(u_0, v_0)$ the ball of radius $\epsilon > 0$ around (u_0, v_0) . For ϵ small enough, this ball is contained in the interior of \widehat{F}_0 . We then have

$$\begin{aligned} \int_{\widehat{F}_0} \operatorname{div} \left(\frac{1}{2} \frac{G_0 \circ d(z_1, z_2)}{(z_1 - u_0)^2 + (z_2 - v_0)^2} \cdot \begin{pmatrix} z_1 - u_0 \\ z_2 - v_0 \end{pmatrix} \right) d\lambda_2(z_1, z_2) \\ = \lim_{\epsilon \rightarrow 0} \int_{\widehat{F}_0 \setminus B_\epsilon(u_0, v_0)} \operatorname{div} \left(\frac{1}{2} \frac{G_0 \circ d(z_1, z_2)}{(z_1 - u_0)^2 + (z_2 - v_0)^2} \cdot \begin{pmatrix} z_1 - u_0 \\ z_2 - v_0 \end{pmatrix} \right) d\lambda_2(z_1, z_2) \\ = \frac{1}{m+2} \cdot \sum_{i=1}^3 \int_{\gamma_i^-}^{\gamma_i^+} \sqrt{w_0^2 + t_i^2 + t^2}^{m+2} \cdot \frac{t_i}{t_i^2 + t^2} dt - \lim_{\epsilon \rightarrow 0} (\pi \cdot G_0(\epsilon^2)). \end{aligned}$$

By definition of G_0 , we now have

$$\pi \cdot G_0(\epsilon^2) = \pi \frac{2}{m+2} \sqrt{w_0^2 + \epsilon^2}^{m+2} \xrightarrow{\epsilon \rightarrow 0} \frac{2\pi}{m+2} |w_0|^{m+2}.$$

Finally, we have for sufficiently small $\epsilon > 0$ that

$$\begin{aligned}
2\pi|w_0|^{m+2} &= \frac{1}{\epsilon} \int_{\partial B_\epsilon(u_0, v_0)} |w_0|^{m+2} ds \\
&= \int_{\partial B_\epsilon(u_0, v_0)} \langle |w_0|^{m+2} \frac{\begin{pmatrix} z_1 - u_0 \\ z_2 - v_0 \end{pmatrix}}{(z_1 - u_0)^2 + (z_2 - v_0)^2}, \frac{\begin{pmatrix} z_1 - u_0 \\ z_2 - v_0 \end{pmatrix}}{\sqrt{(z_1 - u_0)^2 + (z_2 - v_0)^2}} \rangle ds \\
&= \int_{\partial \hat{F}_0} \langle |w_0|^{m+2} \frac{\begin{pmatrix} z_1 - u_0 \\ z_2 - v_0 \end{pmatrix}}{(z_1 - u_0)^2 + (z_2 - v_0)^2}, \hat{m} \rangle ds \\
&= \sum_{i=1}^3 \int_{\gamma_i^-}^{\gamma_i^+} |w_0|^{m+2} \frac{t_i}{t_i^2 + t^2} dt.
\end{aligned}$$

This finishes the proof. \square

Theorem 2.6.10. *Let $w_0 \neq 0$, $t_1 \neq 0$, $t_2 \neq 0$, and $t_3 \neq 0$. For $1 \leq i \leq 3$, let*

$$\beta_i := \arctan\left(\frac{t_i \gamma_i^+}{R_i^2 + |w_0| R_i^+}\right) - \arctan\left(\frac{t_i \gamma_i^-}{R_i^2 + |w_0| R_i^-}\right),$$

and let $\beta := \beta_1 + \beta_2 + \beta_3$. Then we have for $m \in \mathbb{Z}$ with $m \neq -2$

$$\int_F R^m dS = \frac{1}{m+2} \cdot \sum_{i=1}^3 \int_{\gamma_i^-}^{\gamma_i^+} \frac{\sqrt{w_0^2 + t_i^2 + t^2}^{m+3} - |w_0|^{m+3}}{\sqrt{w_0^2 + t_i^2 + t^2}} \cdot \frac{t_i}{t_i^2 + t^2} dt - \frac{|w_0|^{m+2} \beta}{m+2}. \quad (2.65)$$

Proof. By lemma 2.6.9, we have

$$\int_F R^m dS = \frac{1}{m+2} \cdot \sum_{i=1}^3 \int_{\gamma_i^-}^{\gamma_i^+} \left(\sqrt{w_0^2 + t_i^2 + t^2}^{m+2} - \sqrt{w_0^2}^{m+2} \right) \cdot \frac{t_i}{t_i^2 + t^2} dt.$$

Now, let

$$f : \mathbb{R} \rightarrow \mathbb{R}; t \mapsto \arctan\left(\frac{t_i \cdot t}{w_0^2 + t_i^2 + |w_0| \cdot \sqrt{w_0^2 + t_i^2 + t^2}}\right).$$

A tedious but straightforward computation then shows that

$$f'(t) = \left(1 - \frac{|w_0|}{\sqrt{w_0^2 + t_i^2 + t^2}}\right) \frac{t_i}{t_i^2 + t^2}.$$

This implies for $1 \leq i \leq 3$ that

$$\begin{aligned}
&\int_{\gamma_i^-}^{\gamma_i^+} \left(\sqrt{w_0^2 + t_i^2 + t^2}^{m+2} - \sqrt{w_0^2}^{m+2} \right) \cdot \frac{t_i}{t_i^2 + t^2} dt \\
&= \int_{\gamma_i^-}^{\gamma_i^+} \frac{\sqrt{w_0^2 + t_i^2 + t^2}^{m+3} - |w_0|^{m+3}}{\sqrt{w_0^2 + t_i^2 + t^2}} \cdot \frac{t_i}{t_i^2 + t^2} dt - |w_0|^{m+2} \int_{\gamma_i^-}^{\gamma_i^+} \left(1 - \frac{|w_0|}{\sqrt{w_0^2 + t_i^2 + t^2}}\right) \cdot \frac{t_i}{t_i^2 + t^2} dt \\
&= \int_{\gamma_i^-}^{\gamma_i^+} \frac{\sqrt{w_0^2 + t_i^2 + t^2}^{m+3} - |w_0|^{m+3}}{\sqrt{w_0^2 + t_i^2 + t^2}} \cdot \frac{t_i}{t_i^2 + t^2} dt - |w_0|^{m+2} \cdot \beta_i. \quad (2.66)
\end{aligned}$$

Summing over i gives the statement of the theorem. \square

By looking at the integral defining β , one can see that, when interpreted as a function of x_0 , there is a unique continuous extension of β to $\mathbb{R}^3 \setminus F$, and that on $\{x_0 \in \mathbb{R}^3 \setminus F \mid w_0 = 0\}$ we have $\beta = 0$.

We believe that the formulation of theorem 2.6.10 is the “correct” way to compute the integral. If you plug the integral into a symbolic integral solver of your choice, you will see that the generated antiderivatives follow a predictable pattern. However, we will not try to prove a general rule and will focus only on computing the special cases that are relevant to us.

Furthermore, note that, for $w_0 \neq 0$, β_i is still well-defined if $t_1 = 0$, $t_2 = 0$, or $t_3 = 0$.

Corollary 2.6.11. *Let β be defined as in theorem 2.6.10. Furthermore, for $1 \leq i \leq 3$ let*

$$f_i := \begin{cases} \ln \left(\frac{R_i^+ + \gamma_i^+}{R_i^- + \gamma_i^-} \right) & \text{if } \gamma_i^- \geq 0 \\ \ln \left(\frac{R_i^- - \gamma_i^-}{R_i^+ - \gamma_i^+} \right) & \text{if } \gamma_i^- < 0 \end{cases}$$

and

$$R_i^s := \frac{1}{R_i^2} \left(\frac{\gamma_i^+}{R_i^+} - \frac{\gamma_i^-}{R_i^-} \right).$$

For $w_0 \neq 0$ we then have

$$\begin{aligned} \int_F \frac{1}{R} dS &= \sum_{i=1}^3 f_i t_i - |w_0| \beta \\ \int_F \frac{1}{R^3} dS &= \frac{\beta}{|w_0|} \\ \int_F \frac{1}{R^5} dS &= \sum_{i=1}^3 \frac{t_i R_i^s}{3w_0^2} + \frac{\beta}{3|w_0|^3} \end{aligned}$$

Proof. First note that the set $\{x_0 \in \mathbb{R}^3 \setminus F \mid w_0 \neq 0, t_1 \neq 0, t_2 \neq 0, t_3 \neq 0\}$ is dense in the set $\{x_0 \in \mathbb{R}^3 \setminus F \mid w_0 \neq 0\}$. Since the above expressions all depend continuously on x_0 , it thus suffices to show the claims under the additional assumptions that $t_1 \neq 0$, $t_2 \neq 0$, and $t_3 \neq 0$.

If we now define for $1 \leq i \leq 3$

$$\begin{aligned} h_{-1}^i(t) &:= -\frac{t_i}{2} \ln \left(\frac{\sqrt{w_0^2 + t_i^2 + t^2} - t}{\sqrt{w_0^2 + t_i^2 + t^2} + t} \right) \\ h_{-3}^i(t) &= 0 \\ h_{-5}^i(t) &= -\frac{t_i \cdot t}{w_0^2(w_0^2 + t_i^2) \sqrt{w_0^2 + t_i^2 + t^2}} \end{aligned}$$

a straightforward computation reveals that for $m \in \{-5, -3, -1\}$ we have

$$(h_m^i)' = \frac{\sqrt{w_0^2 + t_i^2 + t^2}^{m+3} - |w_0|^{m+3}}{\sqrt{w_0^2 + t_i^2 + t^2}} \cdot \frac{t_i}{t_i^2 + t^2}$$

Using theorem 2.6.10, this now enables us to compute the integrals. This directly yields the claimed formulas for $m = -5$ and $m = -3$. For $m = -1$, note that expanding the argument of

the logarithm with $\sqrt{w_0^2 + t_i^2 + t^2} + t$ (resp. with $\sqrt{w_0^2 + t_i^2 + t^2} - t$) leads to

$$\begin{aligned} h_{-1}^i(t) &= -t_i \cdot \ln \left(\frac{R_i}{\sqrt{w_0^2 + t_i^2 + t^2} + t} \right) \\ &= -t_i \cdot \ln \left(\frac{\sqrt{w_0^2 + t_i^2 + t^2} - t}{R_i} \right) \end{aligned}$$

Using the first expression to evaluate the integrals leads to the first case in the definition of f_i , while using the second expression leads to the second case for f_i . We have thus shown the validity of the formula for $\int_F \frac{1}{R} dS$. \square

Note that, under our assumption that $w_0 \neq 0$, the case distinction in the definition of f_i is not necessary. This case distinction is also not present in the original paper [235] and was introduced by us. The reason for introducing this case distinction will become clear once we actually compute the local subtraction right-hand side integrals. There, we need to generalize the integrals to the case $w_0 = 0$, and we will see that the original formula from [235] can potentially run into a singularity. Concretely, in the case that x_0 lies exactly on the ray defined by the edge s_i , we have $R_i^- = -\gamma_i^-$ and $R_i^+ = -\gamma_i^+$, and the first formula is not suited to compute the corresponding integral. The second definition, on the other hand, is still valid in this case. In particular, we see that f_i extends uniquely to a continuous function on $\mathbb{R}^3 \setminus F$, and can be computed in a numerically stable form with the given expression.

Furthermore, note that

$$\begin{aligned} R_i^s &= \frac{1}{R_i^2} \left(\frac{\gamma_i^+}{R_i^+} - \frac{\gamma_i^-}{R_i^-} \right) = \frac{\frac{\gamma_i^+}{R_i^+} - \frac{\gamma_i^-}{R_i^-}}{R_i^2} = \frac{\frac{\gamma_i^+}{R_i^+} - \frac{\gamma_i^-}{R_i^-}}{R_i^2} \cdot \frac{\frac{\gamma_i^+}{R_i^+} + \frac{\gamma_i^-}{R_i^-}}{\frac{\gamma_i^+}{R_i^+} + \frac{\gamma_i^-}{R_i^-}} \\ &= \frac{\frac{(\gamma_i^+)^2}{(R_i^+)^2} - \frac{(\gamma_i^-)^2}{(R_i^-)^2}}{R_i^2 \cdot \left(\frac{\gamma_i^+}{R_i^+} + \frac{\gamma_i^-}{R_i^-} \right)} = \frac{\frac{(\gamma_i^+)^2 \cdot (R_i^-)^2 - (\gamma_i^-)^2 \cdot (R_i^+)^2}{(R_i^+ \cdot R_i^-)^2}}{R_i^2 \cdot \left(\frac{\gamma_i^+}{R_i^+} + \frac{\gamma_i^-}{R_i^-} \right)} \\ &= \frac{\frac{R_i^2 \cdot ((\gamma_i^+)^2 - (\gamma_i^-)^2)}{(R_i^+ \cdot R_i^-)^2}}{R_i^2 \cdot \left(\frac{\gamma_i^+}{R_i^+} + \frac{\gamma_i^-}{R_i^-} \right)} = \frac{1}{\left(\frac{\gamma_i^+}{R_i^+} + \frac{\gamma_i^-}{R_i^-} \right)} \cdot \frac{((\gamma_i^+)^2 - (\gamma_i^-)^2)}{(R_i^+ \cdot R_i^-)^2}. \end{aligned}$$

Note that this shows that R_i^s also uniquely extends to a continuous function on $\mathbb{R}^3 \setminus F$, and the last expression on the right can be used to evaluate R_i^s in a numerically stable way on or close to $\{x_0 \in \mathbb{R}^3 \setminus F \mid w_0 = 0\}$.

Thankfully, arriving at corollary 2.6.11 is the most technically involved step in the derivation of the analytical formulas. The remaining ingredients for the analytical computation of the local subtraction right-hand side are integrals of products of the form $R^m \cdot p$, where p is a polynomial of degree 1 or 2 and $m \in \{-1, -3, -5\}$. Their derivations are our next goal.

To simplify the following computations, we follow [75] and introduce a translation

$$\tau : \mathbb{R}^2 \rightarrow \mathbb{R}^2; \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \mapsto \begin{pmatrix} z_1 + u_0 \\ z_2 + v_0 \end{pmatrix}.$$

Define $\widehat{F} := \tau^{-1}(\widehat{F}_0)$. Then $\mu_F^\tau := \mu_F \circ \tau$ is a map of F with gramian determinant 1, and the normals of \widehat{F} are still given by \widehat{m}_1 , \widehat{m}_2 , and \widehat{m}_3 . The reason for introducing this additional translation is that we then have for $(z_1, z_2) \in \widehat{F}$ that

$$R(\mu_F^\tau(z)) = \|\mu_F^\tau(z) - x_0\| = \sqrt{w_0^2 + z_1^2 + z_2^2}.$$

Furthermore, we have by equation (2.64) for $(z_1, z_2) \in \widehat{F}$ that

$$\begin{aligned} \begin{pmatrix} \varphi_1 \\ \varphi_2 \\ \varphi_3 \end{pmatrix} \circ \mu_F^\tau(z_1, z_2) &= \begin{pmatrix} 1 & -\frac{1}{\gamma_3} & \frac{\frac{u_3}{\gamma_3} - 1}{v_3} \\ 0 & \frac{1}{\gamma_3} & -\frac{u_3}{\gamma_3 v_3} \\ 0 & 0 & \frac{1}{v_3} \end{pmatrix} \cdot \begin{pmatrix} 1 \\ z_1 + u_0 \\ z_2 + v_0 \end{pmatrix} \\ &= \underbrace{\begin{pmatrix} 1 - \frac{u_0}{\gamma_3} + \frac{\frac{u_3}{\gamma_3} - 1}{v_3} v_0 & -\frac{1}{\gamma_3} & \frac{\frac{u_3}{\gamma_3} - 1}{v_3} \\ \frac{u_0}{\gamma_3} - \frac{u_3 v_0}{\gamma_3 v_3} & \frac{1}{\gamma_3} & -\frac{u_3}{\gamma_3 v_3} \\ \frac{v_0}{v_3} & 0 & \frac{1}{v_3} \end{pmatrix}}_{=: N \in \mathbb{R}^{3 \times 3}} \cdot \begin{pmatrix} 1 \\ z_1 \\ z_2 \end{pmatrix}. \quad (2.67) \end{aligned}$$

In the following, we will furthermore write

$$z_1 : \widehat{F} \rightarrow \mathbb{R}; (a, b) \mapsto a$$

and

$$z_2 : \widehat{F} \rightarrow \mathbb{R}; (a, b) \mapsto b,$$

and define $e_1 = (1, 0)$ and $e_2 = (0, 1)$. To state the results in a symmetric manner, we also define $u_1 := u$ and $u_2 := v$.

The following two theorems are a combination of results from [235], [75], and [18].

Theorem 2.6.12. *We use the same notation as in corollary 2.6.11. Let $w_0 \neq 0$. Then we have for $1 \leq j \leq 2$*

$$\begin{aligned} \int_{\widehat{F}} \left(\frac{1}{R} \circ \mu_F^\tau \right) \cdot z_j \, d\lambda_2 &= \frac{1}{2} \cdot \sum_{i=1}^3 \langle u_j, m_i \rangle \cdot (\gamma_i^+ R_i^+ - \gamma_i^- R_i^- + R_i^2 f_i) \\ \int_{\widehat{F}} \left(\frac{1}{R^3} \circ \mu_F^\tau \right) \cdot z_j \, d\lambda_2 &= - \sum_{i=1}^3 \langle u_j, m_i \rangle \cdot f_i \\ \int_{\widehat{F}} \left(\frac{1}{R^5} \circ \mu_F^\tau \right) \cdot z_j \, d\lambda_2 &= - \frac{1}{3} \cdot \sum_{i=1}^3 \langle u_j, m_i \rangle \cdot R_i^s. \end{aligned}$$

Proof. Let $m \in \mathbb{Z}$. Then we have

$$R^m \circ \mu_F^\tau = \sqrt{w_0^2 + z_1^2 + z_2^2}^m,$$

and thus

$$\frac{\partial}{\partial z_j} (R^m \circ \mu_F^\tau) = m R^{m-2} \circ \mu_F^\tau \cdot z_j,$$

which in turn implies for $m \neq -2$ that

$$\begin{aligned}
\int_{\widehat{F}} (R^m \circ \mu_F^\tau) \cdot z_j d\lambda_2 &= \int_{\widehat{F}} \operatorname{div} \left(\frac{1}{m+2} \cdot R^{m+2} \cdot e_j \right) d\lambda_2 \\
&= \int_{\partial \widehat{F}} \frac{R^{m+2} \circ \mu_F^\tau}{m+2} \langle e_j, \widehat{m} \rangle ds \\
&= \sum_{i=1}^3 \frac{\langle u_j, m_i \rangle}{m+2} \int_{[p_{i+1}, p_{i+1}]} R^{m+2} ds \\
&= \sum_{i=1}^3 \langle u_j, m_i \rangle \cdot \int_{\gamma_i^-}^{\gamma_i^+} \frac{\sqrt{w_0^2 + t_i^2 + t^2}^{m+2}}{m+2} dt.
\end{aligned}$$

For $m = -1$, define

$$H_{-1}^i(t) := \frac{1}{2} t \sqrt{w_0^2 + t_i^2 + t^2} - \frac{w_0^2 + t_i^2}{4} \ln \left(\frac{\sqrt{w_0^2 + t_i^2 + t^2} - t}{\sqrt{w_0^2 + t_i^2 + t^2} + t} \right).$$

Utilizing previous computations from the $m = -1$ case in corollary 2.6.11, one can quickly check that $(H_{-1}^i)'(t) = \sqrt{w_0^2 + t_i^2 + t^2}$. This directly yields the claimed expression for the integral $\int_{\widehat{F}} (\frac{1}{R} \circ \mu_F^\tau) \cdot z_j d\lambda_2$. The computation for the $m = -1$ case in corollary 2.6.11 immediately yields the claimed expression for $\int_{\widehat{F}} (\frac{1}{R^3} \circ \mu_F^\tau) \cdot z_j d\lambda_2$. Finally, the computations from corollary 2.6.11 for the $m = -5$ case also immediately yield the claimed expression for $\int_{\widehat{F}} (\frac{1}{R^5} \circ \mu_F^\tau) \cdot z_j d\lambda_2$. \square

In the following, we will, for the sake of comprehensibility, write ∇_2 for the gradient of a function $h : \mathbb{R}^2 \rightarrow \mathbb{R}$, and ∇_3 for the gradient of a function $g : \mathbb{R}^3 \rightarrow \mathbb{R}$. Furthermore, we want to remind the reader that the divergence theorem implies that, for a sufficiently smooth $f : \mathbb{R}^n \rightarrow \mathbb{R}$ and a regular enough region $\Omega \subset \mathbb{R}^n$, we have

$$\int_{\Omega} \nabla f d\lambda_n = \int_{\partial \Omega} f \cdot \eta dS \in \mathbb{R}^n,$$

where η denotes the outer unit normal of Ω .

Theorem 2.6.13. *We use the same notation as in corollary 2.6.11. Let $w_0 \neq 0$. For $1 \leq i \leq 3$ let $\widehat{s}_i := \frac{q_{i+2} - q_{i+1}}{\|q_{i+2} - q_{i+1}\|}$. Then we have for $1 \leq j \leq 2$*

$$\begin{aligned}
\int_{\widehat{F}} \nabla_2 \left(\frac{1}{R} \circ \mu_F^\tau \right) d\lambda_2 &= \sum_{i=1}^3 f_i \cdot \widehat{m}_i \\
\int_{\widehat{F}} \nabla_2 \left(\frac{1}{R^3} \circ \mu_F^\tau \right) d\lambda_2 &= \sum_{i=1}^3 R_i^s \cdot \widehat{m}_i \\
\int_{\widehat{F}} \nabla_2 \left(\frac{1}{R} \circ \mu_F^\tau \right) \cdot z_j d\lambda_2 &= |w_0| \cdot \beta \cdot e_j + \sum_{i=1}^3 \langle u_j, s_i \rangle \cdot ((R_i^+ - R_i^-) \cdot \widehat{m}_i - t_i \cdot f_i \cdot \widehat{s}_i) \\
\int_{\widehat{F}} \nabla_2 \left(\frac{1}{R^3} \circ \mu_F^\tau \right) \cdot z_j d\lambda_2 &= \sum_{i=1}^3 \langle u_j, \left(\frac{1}{R_i^-} - \frac{1}{R_i^+} \right) \cdot s_i + t_i \cdot R_i^s \cdot m_i \rangle \cdot \widehat{m}_i - \frac{\beta}{|w_0|} \cdot e_j
\end{aligned}$$

Proof. For $m \in \mathbb{Z}$ we have

$$\int_{\widehat{F}} \nabla_2 (R^m \circ \mu_F^\tau) d\lambda_2 = \sum_{i=1}^3 \left(\int_{\gamma_i^-}^{\gamma_i^+} \sqrt{w_0^2 + t_i^2 + t^2}^m dt \right) \cdot \widehat{m}_i.$$

The computations in the proof of corollary 2.6.11 then directly imply the stated formulas for $m = -1$ and $m = -3$.

Since for scalar fields φ and ψ we have $\nabla(\varphi \cdot \psi) = \psi \nabla \varphi + \varphi \nabla \psi$, we have for an arbitrary differentiable scalar field φ on \widehat{F} and $m \in \mathbb{Z}$ that

$$\varphi \cdot \nabla_2 (R^m \circ \mu_F^\tau) = \nabla_2 (\varphi \cdot R^m \circ \mu_F^\tau) - (R^m \circ \mu_F^\tau) \cdot \nabla \varphi,$$

and thus

$$\begin{aligned} & \int_{\widehat{F}} \varphi \cdot \nabla_2 (R^m \circ \mu_F^\tau) d\lambda_2 \\ &= \sum_{i=1}^3 \int_{\gamma_i^-}^{\gamma_i^+} \sqrt{w_0^2 + t_i^2 + t^2}^m \cdot \varphi \circ (\mu_F^\tau)^{-1} (\rho + t_i \cdot m_i + t \cdot s_i) dt \cdot \widehat{m}_i - \int_{\widehat{F}} (R^m \circ \mu_F^\tau) \cdot \nabla \varphi d\lambda_2. \end{aligned}$$

A direct computation furthermore yields $(\mu_F^\tau)^{-1}(x) = \begin{pmatrix} \langle u_1, x - \rho \rangle \\ \langle u_2, x - \rho \rangle \end{pmatrix}$ (remember that $u_1 := u$ and $u_2 := v$). This implies for $\varphi = z_j$ and $1 \leq i \leq 3$ that

$$\varphi \circ (\mu_F^\tau)^{-1} (\rho + t_i \cdot m_i + t \cdot s_i) = \langle u_j, m_i \rangle \cdot t_i + \langle u_j, s_i \rangle \cdot t.$$

Let $m = -1$. By corollary 2.6.11 in combination with $e_j = \langle e_j, \widehat{s}_i \rangle \cdot \widehat{s}_i + \langle e_j, \widehat{m}_i \rangle \cdot \widehat{m}_i$, and utilizing the already computed integrals, we can then compute

$$\begin{aligned} & \int_{\widehat{F}} \nabla_2 \left(\frac{1}{R} \circ \mu_F^\tau \right) \cdot z_j d\lambda_2 \\ &= \sum_{i=1}^3 \int_{\gamma_i^-}^{\gamma_i^+} \frac{\langle u_j, m_i \rangle \cdot t_i}{\sqrt{w_0^2 + t_i^2 + t^2}} + \frac{\langle u_j, s_i \rangle \cdot t}{\sqrt{w_0^2 + t_i^2 + t^2}} dt \cdot \widehat{m}_i - \int_F \frac{1}{R} dS \cdot e_j \\ &= \sum_{i=1}^3 \left(\langle u_j, m_i \rangle \cdot t_i f_i + \langle u_j, s_i \rangle \left[\sqrt{w_0^2 + t_i^2 + t^2} \right]_{\gamma_i^-}^{\gamma_i^+} \right) \cdot \widehat{m}_i - \left(\sum_{i=1}^3 f_i t_i - |w_0| \beta \right) \cdot e_j \\ & \quad \begin{matrix} \langle u_j, m_i \rangle = \langle e_j, \widehat{m}_i \rangle \\ \langle u_j, s_i \rangle = \langle e_j, \widehat{s}_i \rangle \end{matrix} |w_0| \beta e_j + \sum_{i=1}^3 \langle u_j, s_i \rangle \cdot ((R_i^+ - R_i^-) \cdot \widehat{m}_i - t_i f_i \widehat{s}_i). \end{aligned}$$

For $m = -3$, we can similarly compute

$$\begin{aligned}
& \int_{\widehat{F}} \nabla_2 \left(\frac{1}{R^3} \circ \mu_F^\tau \right) \cdot z_j d\lambda_2 \\
&= \sum_{i=1}^3 \int_{\gamma_i^-}^{\gamma_i^+} \frac{\langle u_j, m_i \rangle \cdot t_i}{\sqrt{w_0^2 + t_i^2 + t^2}^3} + \frac{\langle u_j, s_i \rangle \cdot t}{\sqrt{w_0^2 + t_i^2 + t^2}^3} dt \cdot \widehat{m}_i - \int_F \frac{1}{R^3} dS \cdot e_j \\
&= \sum_{i=1}^3 \left(\langle u_j, m_i \rangle \cdot t_i \cdot R_i^s + \langle u_j, s_i \rangle \cdot \left[-\frac{1}{\sqrt{w_0^2 + t_i^2 + t^2}} \right]_{\gamma_i^-}^{\gamma_i^+} \right) \cdot \widehat{m}_i - \int_F \frac{1}{R^3} dS \cdot e_j \\
&= \sum_{i=1}^3 \left(\langle u_j, m_i \rangle \cdot t_i \cdot R_i^s + \langle u_j, s_i \rangle \cdot \left(\frac{1}{R_i^-} - \frac{1}{R_i^+} \right) \right) \cdot \widehat{m}_i - \frac{\beta}{|w_0|} \cdot e_j.
\end{aligned}$$

□

The final theorem in this subsection is due to us. It concerns integrals against quadratic polynomials, which arise when computing the local subtraction right-hand side in a CEM approach.

Theorem 2.6.14. *We use the same notation as in corollary 2.6.11 and theorem 2.6.13. Let $w_0 \neq 0$. We then have for $1 \leq j \leq 2$*

$$\int_{\widehat{F}} \left(\frac{1}{R^3} \circ \mu_F^\tau \right) \cdot z_j^2 d\lambda_2 = -|w_0| \cdot \beta + \sum_{i=1}^3 t_i \cdot f_i - \langle u_j, m_i \rangle \cdot (\langle u_j, s_i \rangle \cdot (R_i^+ - R_i^-) + \langle u_j, m_i \rangle \cdot t_i \cdot f_i)$$

$$\int_{\widehat{F}} \left(\frac{1}{R^3} \circ \mu_F^\tau \right) \cdot z_1 \cdot z_2 d\lambda_2 = - \sum_{i=1}^3 \langle u_1, m_i \rangle \cdot (\langle u_2, m_i \rangle \cdot t_i \cdot f_i + \langle u_2, s_i \rangle (R_i^+ - R_i^-))$$

Furthermore, we have

$$\begin{aligned}
& \int_{\widehat{F}} \nabla_2 \left(\frac{1}{R} \circ \mu_F^\tau \right) \cdot z_j^2 d\lambda_2 = \\
& \sum_{i=1}^3 \left(\langle u_j, m_i \rangle^2 \cdot t_i^2 \cdot f_i + 2 \cdot \langle u_j, m_i \rangle \cdot \langle u_j, s_i \rangle \cdot t_i \cdot (R_i^+ - R_i^-) + \frac{\langle u_j, s_i \rangle^2}{2} \cdot (\gamma_i^+ \cdot R_i^+ - \gamma_i^- \cdot R_i^- - R_i^2 \cdot f_i) \right) \widehat{m}_i \\
& \quad - \left(\sum_{i=1}^3 \langle u_j, m_i \rangle \cdot (\gamma_i^+ \cdot R_i^+ - \gamma_i^- \cdot R_i^- + R_i^2 \cdot f_i) \right) \cdot e_j
\end{aligned}$$

$$\begin{aligned}
& \int_{\widehat{F}} \nabla_2 \left(\frac{1}{R} \circ \mu_F^\tau \right) \cdot z_1 \cdot z_2 d\lambda_2 = \\
& \sum_{i=1}^3 \left(\langle u_1, m_i \rangle \cdot \langle u_2, m_i \rangle \cdot t_i^2 \cdot f_i + (\langle u_1, m_i \rangle \cdot \langle u_2, s_i \rangle + \langle u_1, s_i \rangle \cdot \langle u_2, m_i \rangle) \cdot t_i \cdot (R_i^+ - R_i^-) \right. \\
& \quad \left. + \frac{\langle u_1, s_i \rangle \cdot \langle u_2, s_i \rangle}{2} \cdot (\gamma_i^+ \cdot R_i^+ - \gamma_i^- \cdot R_i^- - R_i^2 \cdot f_i) \right) \cdot \widehat{m}_i \\
& \quad - \frac{1}{2} \cdot \sum_{i=1}^3 (\gamma_i^+ \cdot R_i^+ - \gamma_i^- \cdot R_i^- + R_i^2 \cdot f_i) \cdot \left(\frac{\langle u_2, m_i \rangle}{\langle u_1, m_i \rangle} \right).
\end{aligned}$$

Proof. First, note that for $1 \leq j \leq 2$ we have

$$\begin{aligned} \frac{\partial}{\partial z_j} \left(z_j \cdot \frac{1}{R} \circ \mu_F^\tau \right) &= \frac{1}{R} \circ \mu_F^\tau + z_j \cdot \left((-1) \cdot \frac{1}{R^3} \circ \mu_F^\tau \cdot z_j \right) \\ &\iff \frac{1}{R^3} \circ \mu_F^\tau \cdot z_j^2 = \frac{1}{R} \circ \mu_F^\tau - \operatorname{div} \left(z_j \cdot \frac{1}{R} \circ \mu_F^\tau \cdot e_j \right). \end{aligned}$$

The integral $\int_F \frac{1}{R} dS$ was already computed in corollary 2.6.11. Furthermore, we have

$$\int_{\widehat{F}} \operatorname{div} \left(z_j \cdot \frac{1}{R} \circ \mu_F^\tau \cdot e_j \right) d\lambda_2 = \sum_{i=1}^3 \langle u_j, m_i \rangle \int_{\gamma_i^-}^{\gamma_i^+} \frac{\langle u_j, m_i \rangle \cdot t_i + \langle u_j, s_i \rangle \cdot t}{\sqrt{w_0^2 + t_i^2 + t^2}} dt$$

Now note that the last integral on the right was already computed in the proof of theorem 2.6.13. Combining the previous results leads to the stated formula for the integral $\int_{\widehat{F}} \left(\frac{1}{R^3} \circ \mu_F^\tau \right) \cdot z_j^2 d\lambda_2$.

Furthermore, note that

$$\operatorname{div} \left(z_2 \cdot \frac{1}{R} \circ \mu_F^\tau \cdot e_1 \right) = -z_1 \cdot z_2 \cdot \frac{1}{R^3} \circ \mu_F^\tau,$$

which implies

$$\int_{\widehat{F}} \left(\frac{1}{R^3} \circ \mu_F^\tau \right) \cdot z_1 \cdot z_2 d\lambda_2 = - \sum_{i=1}^3 \langle u_1, m_i \rangle \cdot \int_{\gamma_i^-}^{\gamma_i^+} \frac{\langle u_2, m_i \rangle \cdot t_i + \langle u_2, s_i \rangle \cdot t}{\sqrt{w_0^2 + t_i^2 + t^2}} dt,$$

which directly yields the stated formula for $\int_{\widehat{F}} \left(\frac{1}{R^3} \circ \mu_F^\tau \right) \cdot z_1 \cdot z_2 d\lambda_2$.

Now, applying the rule from the proof of theorem 2.6.13 for $\varphi = z_j^2$ and $m = -1$ implies that

$$\begin{aligned} \int_{\widehat{F}} \nabla_2 \left(\frac{1}{R} \circ \mu_F^\tau \right) \cdot z_j^2 d\lambda_2 &= \\ &= \sum_{i=1}^3 \int_{\gamma_i^-}^{\gamma_i^+} \frac{(\langle u_j, m_i \rangle \cdot t_i + \langle u_j, s_i \rangle \cdot t)^2}{\sqrt{w_0^2 + t_i^2 + t^2}} dt \cdot \widehat{m}_i - 2 \cdot \int_{\widehat{F}} \frac{1}{R} \circ \mu_F^\tau \cdot z_j d\lambda_2 \cdot e_j. \end{aligned}$$

The last integral on the right-hand side can be evaluated using theorem 2.6.12. The first integral on the left-hand side can be evaluated by expanding the square. The constant and linear terms in t can be handled using previously computed integrals. The new integral appearing is $\int_{\gamma_i^-}^{\gamma_i^+} \frac{t^2}{\sqrt{w_0^2 + t_i^2 + t^2}} dt$. Defining

$$v(t) := \frac{1}{2} \cdot t \cdot \sqrt{w_0^2 + t_i^2 + t^2} - \frac{w_0^2 + t_i^2}{4} \cdot \ln \left(\frac{\sqrt{w_0^2 + t_i^2 + t^2} + t}{\sqrt{w_0^2 + t_i^2 + t^2} - t} \right),$$

a straightforward computation shows that $v'(t) = \frac{t^2}{\sqrt{w_0^2 + t_i^2 + t^2}}$ (Note the subtle difference compared to the definition of H_{-1}^i in theorem 2.6.12). We thus have

$$\int_{\gamma_i^-}^{\gamma_i^+} \frac{t^2}{\sqrt{w_0^2 + t_i^2 + t^2}} dt = \frac{1}{2} \cdot (\gamma_i^+ \cdot R_i^+ - \gamma_i^- \cdot R_i^- - R_i^2 \cdot f_i).$$

Carefully gathering all terms yields the claimed expression for the integral $\int_{\widehat{F}} \nabla_2 \left(\frac{1}{R} \circ \mu_F^\tau \right) \cdot z_j^2 d\lambda_2$. Finally, applying the rule from the proof of theorem 2.6.13 for $\varphi = z_1 \cdot z_2$ and $m = -1$ implies that

$$\begin{aligned} & \int_{\widehat{F}} \nabla_2 \left(\frac{1}{R} \circ \mu_F^\tau \right) \cdot z_1 \cdot z_2 d\lambda_2 \\ &= \sum_{i=1}^3 \int_{\gamma_i^-}^{\gamma_i^+} \frac{(\langle u_1, m_i \rangle \cdot t_i + \langle u_1, s_i \rangle \cdot t) \cdot (\langle u_2, m_i \rangle \cdot t_i + \langle u_2, s_i \rangle \cdot t)}{\sqrt{w_0^2 + t_i^2 + t^2}} dt \cdot \widehat{m}_i \\ & \quad - \int_{\widehat{F}} \frac{1}{R} \circ \mu_F^\tau \cdot (z_2 \cdot e_1 + z_1 \cdot e_2) d\lambda_2 \end{aligned}$$

Multiplying out all products and gathering by powers of t , one sees that all integrals that arise have already been computed before in this subsection. This yields the analytical expression for $\int_{\widehat{F}} \nabla_2 \left(\frac{1}{R} \circ \mu_F^\tau \right) \cdot z_1 \cdot z_2 d\lambda_2$. \square

We now finally have enough recipes in our “cookbook” to derive the analytical expressions for the isotropic local subtraction right-hand side.

Computing the isotropic local subtraction right-hand side

In the following, we will assume that the local subtraction cutoff function χ is affine on each mesh element. In our implementation, as we will present it in section 2.7.1, this is always the case for tetrahedral meshes with piecewise affine trial functions, as we choose χ as a FEM trial function.

Let $\sigma^\infty \in \mathbb{R}$ denote an isotropic conductivity value. As derived in theorem 2.6.2, the corresponding dipole potential in a homogeneous unbounded volume conductor is then given by

$$u^\infty(x) = \frac{1}{4\pi\sigma^\infty} \frac{\langle M, x - x_0 \rangle}{\|x - x_0\|^3}.$$

In the following, let $\vec{R}(x) := x - x_0$, and $R(x) = \|\vec{R}(x)\|$. Note that

$$\frac{\vec{R}}{R^3} = -\nabla_3 \left(\frac{1}{R} \right).$$

Furthermore, let $F \subset \mathbb{R}^3$ denote a triangle. We use the notation from the preceding sections. First, we will investigate how to reduce integrals over F to integrals over \widehat{F} . This result is due to [75].

Lemma 2.6.15. *Let $U \subset \mathbb{R}^3$ denote an open environment of F . Let $f : U \rightarrow \mathbb{R}$ be differentiable. Then*

$$\nabla_3(f) \circ \mu_F^\tau = (u, v) \cdot \nabla_2(f \circ \mu_F^\tau) + \frac{\partial f}{\partial w} \circ \mu_F^\tau \cdot w,$$

where $\frac{\partial f}{\partial w}$ denotes the directional derivative of f in direction w .

Proof. By the chain rule, we have

$$\nabla_2(f \circ \mu_F^\tau) = J_{\mu_F^\tau}^\top \cdot \nabla_3(f) \circ \mu_F^\tau = (u, v)^\top \cdot \nabla_3(f) \circ \mu_F^\tau.$$

Now, since $\{u, v, w\}$ is an orthogonal basis of \mathbb{R}^3 we thus have

$$\begin{aligned}\nabla_3(f) \circ \mu_F^\tau &= (u, v, w) \cdot (u, v, w)^\top \cdot \nabla_3(f) \circ \mu_F^\tau \\ &= (u, v) \cdot \nabla_2(f \circ \mu_F^\tau) + \langle w, \nabla_3(f) \rangle \circ \mu_F^\tau \cdot w.\end{aligned}$$

Since $\langle w, \nabla_3(f) \rangle = \frac{\partial f}{\partial w}$, this finishes the proof. \square

By noticing that for $x \in F$ we have $\langle w, \vec{R}(x) \rangle = w_0$, we in particular get for $f = u^\infty$ that

$$\begin{aligned}u^\infty \circ \mu_F^\tau &= -\frac{1}{4\pi\sigma^\infty} \langle M, \nabla_3 \left(\frac{1}{R} \right) \circ \mu_F^\tau \rangle \\ &= \frac{1}{4\pi\sigma^\infty} \langle M, w_0 \cdot \left(\frac{1}{R^3} \circ \mu_F^\tau \right) \cdot w - (u, v) \cdot \nabla_2 \left(\frac{1}{R} \circ \mu_F^\tau \right) \rangle.\end{aligned}\quad (2.68)$$

Furthermore, for an arbitrary $a \in \mathbb{R}^3$ we have

$$\nabla_3 \left(\frac{\langle a, \vec{R} \rangle}{R^3} \right) = \frac{1}{R^3} \cdot a - 3 \cdot \frac{\langle a, \vec{R} \rangle}{R^5} \cdot \vec{R}.$$

In combination with lemma 2.6.15, this yields

$$\nabla_3 \left(\frac{\langle w, \vec{R} \rangle}{R^3} \right) \circ \mu_F^\tau = w_0 \cdot (u, v) \cdot \nabla_2 \left(\frac{1}{R^3} \circ \mu_F^\tau \right) + \left(\frac{1}{R^3} - 3 \cdot \frac{w_0^2}{R^5} \right) \cdot w.\quad (2.69)$$

We need one final lemma before deriving the analytical expressions.

Lemma 2.6.16. *Let $F : U \rightarrow \mathbb{R}^n$, $U \subset \mathbb{R}^n$ open, be differentiable. If F is closed, i.e. $\frac{\partial F_i}{\partial x_j} = \frac{\partial F_j}{\partial x_i}$ for all $1 \leq i, j \leq n$, and $a, b \in \mathbb{R}^n$ are arbitrary, we have*

$$\langle a, \nabla \langle b, F \rangle \rangle = \langle b, \nabla \langle a, F \rangle \rangle.$$

Proof. We have

$$\langle a, \nabla \langle b, F \rangle \rangle = \sum_{i=1}^n a_i \cdot \frac{\partial}{\partial x_i} \sum_{j=1}^n b_j F_j = \sum_{i,j=1}^n a_i b_j \frac{\partial F_j}{\partial x_i} = \sum_{i,j=1}^n a_i b_j \frac{\partial F_i}{\partial x_j} = \langle b, \nabla \langle a, F \rangle \rangle.$$

\square

Note that all gradient fields are closed, i.e., we can apply lemma 2.6.16 to the field $\frac{\vec{R}}{R^3} = \nabla(-\frac{1}{R})$. We are now ready to compute the local subtraction right-hand sides. We start with the surface integrals. These integrals were first computed by Beltrachini in [18].

Theorem 2.6.17 (Local subtraction surface integrals). *Let $\varphi := \begin{pmatrix} \varphi_1 \\ \varphi_2 \\ \varphi_3 \end{pmatrix}$, where φ_1 , φ_2 , and φ_3*

denote the affine test functions on the triangle F , as defined directly before (2.64). Let $N \in \mathbb{R}^{3 \times 3}$ be defined as in 2.67. Using the notation from the preceding paragraphs, we define

$$I_0 := \sum_{i=1}^3 R_i^s \cdot (w_0 \cdot m_i - t_i \cdot w)$$

and for $1 \leq j \leq 2$

$$I_j := -\operatorname{sign}(w_0) \cdot \beta \cdot u_j + w_0 \cdot \sum_{i=1}^3 \langle u_j, \left(\frac{1}{R_i^-} - \frac{1}{R_i^+} \right) \cdot s_i + t_i \cdot R_i^s \cdot m_i \rangle \cdot m_i \\ + \left(\sum_{i=1}^3 \langle u_j, m_i \rangle (w_0^2 \cdot R_i^s - f_i) \right) \cdot w$$

Then we have

$$\int_F \langle \sigma^\infty \nabla u^\infty, w \rangle \varphi dS = \frac{N}{4\pi} \cdot \begin{pmatrix} \langle I_0, M \rangle \\ \langle I_1, M \rangle \\ \langle I_2, M \rangle \end{pmatrix}.$$

Furthermore, a unit outer normal with respect to the face F is given by either w or $-w$, depending on the tetrahedron whose face is given by F .

Proof. Note that $\int_F \langle \sigma^\infty \nabla u^\infty, w \rangle \varphi dS$, N , I_0 , I_1 , and I_2 all depend continuously on the dipole position $x_0 \in \mathbb{R}^3 \setminus \bar{F}$. We can hence WLOG assume that $w_0 \neq 0$

We now have

$$\begin{pmatrix} \int_F \langle \sigma^\infty \nabla u^\infty, \eta \rangle \varphi_1 dS \\ \int_F \langle \sigma^\infty \nabla u^\infty, \eta \rangle \varphi_2 dS \\ \int_F \langle \sigma^\infty \nabla u^\infty, \eta \rangle \varphi_3 dS \end{pmatrix} \\ = \int_F \langle \sigma^\infty \nabla u^\infty, \eta \rangle \begin{pmatrix} \varphi_1 \\ \varphi_2 \\ \varphi_3 \end{pmatrix} dS = \frac{1}{4\pi} \cdot \int_F \langle \nabla_3 \langle M, \frac{\vec{R}}{R^3} \rangle, w \rangle \cdot \begin{pmatrix} \varphi_1 \\ \varphi_2 \\ \varphi_3 \end{pmatrix} dS \\ \stackrel{\text{Lemma 2.6.16}}{=} \frac{1}{4\pi} \cdot \int_F \langle \nabla_3 \langle w, \frac{\vec{R}}{R^3} \rangle, M \rangle \cdot \begin{pmatrix} \varphi_1 \\ \varphi_2 \\ \varphi_3 \end{pmatrix} dS \\ = \frac{N}{4\pi} \cdot \begin{pmatrix} \langle \int_{\hat{F}} \nabla_3 \langle w, \frac{\vec{R}}{R^3} \rangle d\lambda_2, M \rangle \\ \langle \int_{\hat{F}} z_1 \cdot \nabla_3 \langle w, \frac{\vec{R}}{R^3} \rangle d\lambda_2, M \rangle \\ \langle \int_{\hat{F}} z_2 \cdot \nabla_3 \langle w, \frac{\vec{R}}{R^3} \rangle d\lambda_2, M \rangle \end{pmatrix}$$

Using corollary 2.6.11 and theorem 2.6.13, we now have

$$\int_{\hat{F}} \nabla_3 \langle w, \frac{\vec{R}}{R^3} \rangle d\lambda_2 = \int_{\hat{F}} w_0 \cdot (u, v) \cdot \nabla_2 \left(\frac{1}{R^3} \circ \mu_F \right) + \left(\frac{1}{R^3} - 3 \cdot \frac{w_0^2}{R^5} \right) \cdot w d\lambda_2 \\ = \sum_{i=1}^3 w_0 \cdot R_i^s \cdot m_i + \frac{\beta}{|w_0|} - \frac{\beta}{|w_0|} - \sum_{i=1}^3 t_i \cdot R_i^s \cdot w = \sum_{i=1}^3 R_i^s \cdot (w_0 \cdot m_i - t_i \cdot w) = I_0.$$

Furthermore, using theorems 2.6.12 and 2.6.13 we have for $1 \leq j \leq 2$

$$\begin{aligned} \int_{\widehat{F}} z_j \cdot \nabla_3 \langle w, \frac{\vec{R}}{R^3} \rangle d\lambda_2 &= \int_{\widehat{F}} w_0 \cdot (u, v) \cdot z_j \cdot \nabla_2 \left(\frac{1}{R^3} \circ \mu_F \right) + z_j \cdot \left(\frac{1}{R^3} - 3 \cdot \frac{w_0^2}{R^5} \right) \cdot w d\lambda_2 \\ &= -\text{sign}(w_0) \cdot \beta \cdot u_j + w_0 \cdot \sum_{i=1}^3 \langle u_j, \left(\frac{1}{R_i^-} - \frac{1}{R_i^+} \right) \cdot s_i + t_i \cdot R_i^s \cdot m_i \rangle \cdot m_i \\ &\quad + \left(\sum_{i=1}^3 \langle u_j, m_i \rangle \cdot (w_0^2 \cdot R_i^s - f_i) \right) \cdot w = I_j. \end{aligned}$$

□

We continue with the patch integral. This integral was also computed for the first time by Beltrachini in [18]. Let $T \subset \mathbb{R}^3$ denote a tetrahedron with triangular faces F_0, F_1, F_2 , and F_3 . Let ψ_i denote the unique affine function that is 1 on the i -th vertex of the tetrahedron and zero on the remaining vertices. We then have

$$\begin{aligned} \begin{pmatrix} \int_T \langle \sigma^c \nabla u^\infty, \nabla \psi_0 \rangle dV \\ \int_T \langle \sigma^c \nabla u^\infty, \nabla \psi_1 \rangle dV \\ \int_T \langle \sigma^c \nabla u^\infty, \nabla \psi_2 \rangle dV \\ \int_T \langle \sigma^c \nabla u^\infty, \nabla \psi_3 \rangle dV \end{pmatrix} &= (\nabla \psi_0, \nabla \psi_1, \nabla \psi_2, \nabla \psi_3)^\top \cdot \sigma^c \cdot \int_T \nabla u^\infty dV \\ &= (\nabla \psi_0, \nabla \psi_1, \nabla \psi_2, \nabla \psi_3)^\top \cdot \sigma^c \cdot \sum_{i=0}^3 \left(\int_{F_i} u^\infty dS \right) \cdot \eta_i, \end{aligned}$$

where η_i denotes the outer unit normal of the face F_i . We thus see that the computation of the patch integral comes down to the computation of $\int_F u^\infty dS$ for a triangle $F \subset \mathbb{R}^3$.

Theorem 2.6.18 (Local subtraction patch integrals). *We use the notation from the preceding paragraphs. We then have*

$$\int_F u^\infty dS = \frac{1}{4\pi\sigma^\infty} \langle M, \text{sign}(w_0) \cdot \beta \cdot w - \sum_{i=1}^3 f_i m_i \rangle.$$

Proof. Note that both sides of the equation above again define continuous functions in x_0 on $\mathbb{R}^3 \setminus F$. We can thus WLOG assume $w_0 \neq 0$. We then have

$$\begin{aligned} \int_F u^\infty dS &= \int_{\widehat{F}} u^\infty \circ \mu_F^\tau d\lambda_2 \stackrel{\text{Eq. 2.68}}{=} \frac{1}{4\pi\sigma^\infty} \langle M, w_0 \cdot \int_F \frac{1}{R^3} dS \cdot w - (u, v) \cdot \int_{\widehat{F}} \nabla_2 \left(\frac{1}{R} \circ \mu_F^\tau \right) d\lambda_2 \rangle \\ &\stackrel{\text{Cor. 2.6.11 and Thm. 2.6.13}}{=} \frac{1}{4\pi\sigma^\infty} \langle M, \text{sign}(w_0) \cdot \beta \cdot w - \sum_{i=1}^3 f_i m_i \rangle \end{aligned}$$

□

Note that [18] makes a sign error in his derivation of theorem 2.6.18, where he confuses the definition of ∇ and ∇' in [75].

We now want to compute the transition integrals. These integrals were computed by us in [91]. Similarly to the patch integral case, we can, for a tetrahedron T , compute

$$\begin{aligned} \begin{pmatrix} \int_T \langle \sigma \nabla (\chi \cdot u^\infty), \nabla \psi_0 \rangle dV \\ \int_T \langle \sigma \nabla (\chi \cdot u^\infty), \nabla \psi_1 \rangle dV \\ \int_T \langle \sigma \nabla (\chi \cdot u^\infty), \nabla \psi_2 \rangle dV \\ \int_T \langle \sigma \nabla (\chi \cdot u^\infty), \nabla \psi_3 \rangle dV \end{pmatrix} &= (\nabla \psi_0, \nabla \psi_1, \nabla \psi_2, \nabla \psi_3)^\top \cdot \sigma \cdot \int_T \nabla (\chi \cdot u^\infty) dV \\ &= (\nabla \psi_0, \nabla \psi_1, \nabla \psi_2, \nabla \psi_3)^\top \cdot \sigma \cdot \sum_{i=0}^3 \left(\int_{F_i} \chi \cdot u^\infty dS \right) \cdot \eta_i. \end{aligned}$$

We thus see that the computation of the transition integral comes down to the computation of $\int_F \chi \cdot u^\infty dS$ for a triangle $F \subset \mathbb{R}^3$. Furthermore, note that this is exactly the electrode DOF integral 2.55. If we manage to derive analytical expressions for this integral, we thus solve two of the five local subtraction integrals in one step.

First note that by construction χ is an affine function. If then p_1, p_2 , and p_3 are the vertices of F , and φ_i is the unique affine function on F that is 1 on p_i and 0 on the remaining vertices, we have

$$\chi = \sum_{i=1}^3 \chi(p_i) \cdot \varphi_i.$$

Define $v_\chi := \begin{pmatrix} \chi(p_1) \\ \chi(p_2) \\ \chi(p_3) \end{pmatrix} \in \mathbb{R}^3$.

Theorem 2.6.19 (Local subtraction transition (resp. electrode DOF) integrals). *Let $N \in \mathbb{R}^{3 \times 3}$ be defined as in 2.67. We use the notation from 2.6.3. Furthermore, let*

$$J_0 := \text{sign}(w_0) \cdot \beta \cdot w - \sum_{i=1}^3 f_i m_i \in \mathbb{R}^3$$

and for $1 \leq j \leq 2$

$$J_j := -|w_0| \cdot \beta \cdot u_j + \sum_{i=1}^3 \langle u_j, s_i \rangle \cdot (t_i \cdot f_i \cdot s_i - (R_i^+ - R_i^-) \cdot m_i) - \left(\sum_{i=1}^3 \langle u_j, m_i \rangle \cdot f_i \right) \cdot w_0 \cdot w \in \mathbb{R}^3$$

We then have

$$\int_F \chi \cdot u^\infty dS = \frac{1}{4\pi\sigma^\infty} \cdot \langle M, (J_0, J_1, J_2) \cdot N^\top \cdot v_\chi \rangle.$$

Proof. As before, we can WLOG assume $w_0 \neq 0$. We now have

$$\begin{aligned} \int_F \chi \cdot u^\infty dS &= \left\langle \int_F \begin{pmatrix} \varphi_1 \\ \varphi_2 \\ \varphi_3 \end{pmatrix} \cdot u^\infty dS, v_\chi \right\rangle = \left\langle N \cdot \int_{\widehat{F}} \begin{pmatrix} 1 \\ z_1 \\ z_2 \end{pmatrix} \cdot u^\infty \circ \mu_F^\tau d\lambda_2, v_\chi \right\rangle \\ &= \left\langle N \cdot \begin{pmatrix} \int_{\widehat{F}} u^\infty \circ \mu_F^\tau d\lambda_2 \\ \int_{\widehat{F}} u^\infty \circ \mu_F^\tau \cdot z_1 d\lambda_2 \\ \int_{\widehat{F}} u^\infty \circ \mu_F^\tau \cdot z_2 d\lambda_2 \end{pmatrix}, v_\chi \right\rangle = \left\langle \begin{pmatrix} \int_{\widehat{F}} u^\infty \circ \mu_F^\tau d\lambda_2 \\ \int_{\widehat{F}} u^\infty \circ \mu_F^\tau \cdot z_1 d\lambda_2 \\ \int_{\widehat{F}} u^\infty \circ \mu_F^\tau \cdot z_2 d\lambda_2 \end{pmatrix}, N^\top \cdot v_\chi \right\rangle. \end{aligned}$$

By theorem 2.6.18, we have

$$\int_{\widehat{F}} u^\infty \circ \mu_F^\tau d\lambda_2 = \frac{1}{4\pi\sigma^\infty} \langle M, \text{sign}(w_0) \cdot \beta \cdot w - \sum_{i=1}^3 f_i m_i \rangle = \frac{1}{4\pi\sigma^\infty} \cdot J_0^\top \cdot M.$$

Furthermore, we have for $1 \leq j \leq 2$ that

$$\begin{aligned} \int_{\widehat{F}} u^\infty \circ \mu_F^\tau \cdot z_j d\lambda_2 & \stackrel{\text{Eq. 2.68}}{=} \frac{1}{4\pi\sigma^\infty} \langle M, w_0 \cdot \left(\int_F \frac{1}{R^3} \circ \mu_F^\tau \cdot z_j d\lambda_2 \right) \cdot w - (u, v) \cdot \int_{\widehat{F}} \nabla_2 \left(\frac{1}{R} \circ \mu_F^\tau \right) \cdot z_j d\lambda_2 \rangle \\ & \stackrel{\text{Thm. 2.6.12 and Thm. 2.6.13}}{=} \frac{1}{4\pi\sigma^\infty} \langle M, w_0 \cdot \left(- \sum_{i=1}^3 \langle u_j, m_i \rangle \cdot f_i \right) \cdot w \\ & \quad - |w_0| \cdot \beta \cdot u_j - \sum_{i=1}^3 \langle u_j, s_i \rangle \left((R_i^+ - R_i^-) \cdot m_i - t_i \cdot f_i \cdot s_i \right) \rangle \\ & = \frac{1}{4\pi\sigma^\infty} \cdot J_j^\top \cdot M. \end{aligned}$$

This finishes the proof. \square

Now, only the electrode interface integrals are left. This computation is also due to us and will be our final theorem in this subsection.

Theorem 2.6.20 (Local subtraction electrode interface integrals). *Let $N \in \mathbb{R}^{3 \times 3}$ be defined as in 2.67. We use the notation from the preceding paragraphs. Let v_χ , J_0 , J_1 , and $J_2 \in \mathbb{R}^3$ be defined as in theorem 2.6.19. Then, let*

$$I = \begin{pmatrix} \int_{\widehat{F}} u^\infty \circ \mu_F^\tau d\lambda_2 & \int_{\widehat{F}} u^\infty \circ \mu_F^\tau \cdot z_1 d\lambda_2 & \int_{\widehat{F}} u^\infty \circ \mu_F^\tau \cdot z_2 d\lambda_2 \\ \int_{\widehat{F}} u^\infty \circ \mu_F^\tau \cdot z_1 d\lambda_2 & \int_{\widehat{F}} u^\infty \circ \mu_F^\tau \cdot z_1^2 d\lambda_2 & \int_{\widehat{F}} u^\infty \circ \mu_F^\tau \cdot z_1 \cdot z_2 d\lambda_2 \\ \int_{\widehat{F}} u^\infty \circ \mu_F^\tau \cdot z_2 d\lambda_2 & \int_{\widehat{F}} u^\infty \circ \mu_F^\tau \cdot z_1 \cdot z_2 d\lambda_2 & \int_{\widehat{F}} u^\infty \circ \mu_F^\tau \cdot z_2^2 d\lambda_2 \end{pmatrix} \in \mathbb{R}^{3 \times 3}.$$

We will write $I = (I_{ij})_{1 \leq i, j \leq 3}$. If φ_1 , φ_2 , and φ_3 again denote the unique affine functions which are 1 on their corresponding vertex and 0 on the remaining vertices, we have

$$\begin{pmatrix} \int_F \chi \cdot u^\infty \cdot \varphi_1 dS \\ \int_F \chi \cdot u^\infty \cdot \varphi_2 dS \\ \int_F \chi \cdot u^\infty \cdot \varphi_3 dS \end{pmatrix} = N \cdot I \cdot N^\top \cdot v_\chi.$$

Here, I can be computed via

$$\begin{aligned} I_{11} &= \frac{1}{4\pi\sigma^\infty} \cdot \langle M, J_0 \rangle \\ I_{12} &= I_{21} = \frac{1}{4\pi\sigma^\infty} \cdot \langle M, J_1 \rangle \\ I_{13} &= I_{31} = \frac{1}{4\pi\sigma^\infty} \cdot \langle M, J_2 \rangle \end{aligned}$$

and if we define for $1 \leq j \leq 2$

$$\begin{aligned} K_j := & w_0 \cdot \left(-|w_0| \cdot \beta + \sum_{i=1}^3 t_i \cdot f_i - \sum_{i=1}^3 \langle u_j, m_i \rangle (\langle u_j, s_i \rangle \cdot (R_i^+ - R_i^-) + \langle u_j, m_i \rangle \cdot t_i \cdot f_i) \right) \cdot w \\ & - \sum_{i=1}^3 \left(\langle u_j, m_i \rangle^2 \cdot t_i^2 \cdot f_i + 2 \langle u_j, m_i \rangle \cdot \langle u_j, s_i \rangle \cdot t_i \cdot (R_i^+ - R_i^-) + \frac{\langle u_j, s_i \rangle^2}{2} \cdot (\gamma_i^+ \cdot R_i^+ - \gamma_i^- \cdot R_i^- - R_i^2 \cdot f_i) \right) \cdot m_i \\ & + \left(\sum_{i=1}^3 \langle u_j, m_i \rangle \cdot (\gamma_i^+ \cdot R_i^+ - \gamma_i^- \cdot R_i^- + R_i^2 \cdot f_i) \right) \cdot u_j \in \mathbb{R}^3 \end{aligned}$$

and

$$\begin{aligned} L := & -w_0 \cdot \left(\sum_{i=1}^3 \langle u_1, m_i \rangle \cdot (\langle u_2, m_i \rangle \cdot t_i \cdot f_i + \langle u_2, s_i \rangle \cdot (R_i^+ - R_i^-)) \right) \cdot w \\ & - \sum_{i=1}^3 \left(\langle u_1, m_i \rangle \cdot \langle u_2, m_i \rangle \cdot t_i^2 \cdot f_i + (\langle u_1, m_i \rangle \cdot \langle u_2, s_i \rangle + \langle u_1, s_i \rangle \cdot \langle u_2, m_i \rangle) \cdot t_i \cdot (R_i^+ - R_i^-) \right. \\ & \quad \left. + \frac{\langle u_1, s_i \rangle \cdot \langle u_2, s_i \rangle}{2} \cdot (\gamma_i^+ \cdot R_i^+ - \gamma_i^- \cdot R_i^- - R_i^2 \cdot f_i) \right) \cdot m_i \\ & + \frac{1}{2} \cdot (u_1, u_2) \cdot \left(\sum_{i=1}^3 (\gamma_i^+ \cdot R_i^+ - \gamma_i^- \cdot R_i^- + R_i^2 \cdot f_i) \left(\frac{\langle u_2, m_i \rangle}{\langle u_1, m_i \rangle} \right) \right) \in \mathbb{R}^3, \end{aligned}$$

we have

$$\begin{aligned} I_{22} &= \frac{1}{4\pi\sigma^\infty} \cdot \langle M, K_1 \rangle \\ I_{33} &= \frac{1}{4\pi\sigma^\infty} \cdot \langle M, K_2 \rangle \\ I_{23} = I_{32} &= \frac{1}{4\pi\sigma^\infty} \cdot \langle M, L \rangle. \end{aligned}$$

Proof. Just as before, we can WLOG assume that $w_0 \neq 0$. We then have

$$\begin{aligned} \left(\frac{\int_F \chi \cdot u^\infty \cdot \varphi_1 dS}{\int_F \chi \cdot u^\infty \cdot \varphi_2 dS} \right) &= \left(\int_F u^\infty \cdot \varphi_i \cdot \varphi_j dS \right)_{1 \leq i, j \leq 3} \cdot v_\chi \\ &= \left(\int_{\widehat{F}} u^\infty \circ \mu_F^\tau \cdot \varphi_i \circ \mu_F^\tau \cdot \varphi_j \circ \mu_F^\tau d\lambda_2 \right)_{1 \leq i, j \leq 3} \cdot v_\chi = N \cdot I \cdot N^\top \cdot v_\chi. \end{aligned}$$

The equations for I_{11} , I_{12} , I_{13} , I_{21} , and I_{31} then directly follow from the proof of theorem 2.6.19. The remaining entries of I can then directly be derived by using equation 2.68, i.e.

$$u^\infty \circ \mu_F^\tau = -\frac{1}{4\pi\sigma^\infty} \langle M, \nabla_3 \left(\frac{1}{R} \right) \rangle = \frac{1}{4\pi\sigma^\infty} \langle M, w_0 \cdot \left(\frac{1}{R^3} \circ \mu_F^\tau \right) \cdot w - (u, v) \cdot \nabla_2 \left(\frac{1}{R} \circ \mu_F^\tau \right) \rangle,$$

and applying theorem 2.6.14. \square

We have now derived analytical expressions for all local subtraction right-hand side integrals in tetrahedral models for element-wise affine trial functions and isotropic source element conductivities.

Implementation

All formulas derived in this section have been implemented into DUNEuro. First, we implemented a test where the analytical expressions are validated by comparing them against the results of numerical quadrature.¹⁴ Furthermore, we implemented the analytical expressions into a DUNE local operator,¹⁵ and structured the local subtraction code in such a way that it uses the analytical formulas if their requirements are fulfilled.

2.6.4 Right-hand side analytical expressions in the anisotropic case

Preliminaries

In this subsection, we will show how we can derive analytical expressions for the integrals 2.51 - 2.55 in the more general case of an anisotropic σ^∞ . We will do this by utilizing the fact that we already know how to compute these integrals in the isotropic case. More concretely, if for $z, v \in \mathbb{R}^3$, we define

$$f_{z,v}(y) = \frac{1}{4\pi} \cdot \frac{\langle v, y - z \rangle}{\|y - z\|^3}, \quad (2.70)$$

then, by theorem 2.6.2, we have $\Delta f_{z,v} = \text{div}(v \cdot \delta_z)$, and we know how to compute the integrals $\int_T \langle \sigma^c \nabla f_{z,v}, \nabla \varphi \rangle dV$, $\int_F \langle \nabla f_{z,v}, \eta \rangle \varphi dS$, $\int_T \langle \sigma \nabla \chi \cdot f_{z,v}, \nabla \varphi \rangle dV$, $\int_F \chi \cdot f_{z,v} \cdot \varphi dS$, and $\int_F \chi \cdot f_{z,v} dS$, where φ is an affine test function.

In the following, we will fix a dipole position $x_0 \in \mathbb{R}^3$ and dipole moment $M \in \mathbb{R}^3$. Let then u^∞ denote a corresponding solution of $\text{div}(\sigma^\infty \nabla u^\infty) = \text{div}(M \cdot \delta_{x_0})$ on \mathbb{R}^3 . If $(\sigma^\infty)^{-1} = A \cdot A^\top$ with $\det(A) > 0$, we then have by theorem 2.6.6 that

$$u^\infty = \det(A) \cdot f_{A^\top x_0, A^\top M} \circ \Phi_{A^\top}.$$

This suggests performing a coordinate change with A^\top .

Theorem 2.6.21. *Let $T \subset \mathbb{R}^3$ be a tetrahedron with corners p_0, p_1, p_2 , and p_3 , and let $F \subset \mathbb{R}^3$ be the triangular facet of the tetrahedron with the corners p_0, p_1 , and p_2 . Let $\eta \in \mathbb{R}^3$ denote the unit outer normal of T on F . Then $\Phi_{A^\top}(T)$ is again a tetrahedron with the corners $\Phi_{A^\top}(p_0), \Phi_{A^\top}(p_1), \Phi_{A^\top}(p_2)$, and $\Phi_{A^\top}(p_3)$, and the unit outer normal of the triangular facet of $\Phi_{A^\top}(T)$ with corners $\Phi_{A^\top}(p_0), \Phi_{A^\top}(p_1)$, and $\Phi_{A^\top}(p_2)$ is given by*

$$\eta_{\Phi_{A^\top}(F)} = \frac{A^{-1} \cdot \eta}{\|A^{-1} \cdot \eta\|}.$$

Proof. By definition, the tetrahedron is given by

$$T = \text{conv}(p_0, p_1, p_2, p_3) = \left\{ \sum_{i=0}^3 \lambda_i \cdot p_i \mid 0 \leq \lambda_i, \sum_{i=1}^3 \lambda_i = 1 \right\}.$$

where conv denotes the convex hull. Since Φ_{A^\top} is linear, it follows that $\Phi_{A^\top}(T)$ is the convex hull of $\Phi_{A^\top}(p_0), \Phi_{A^\top}(p_1), \Phi_{A^\top}(p_2)$, and $\Phi_{A^\top}(p_3)$. Since $\det(A^\top) = \det(A) > 0$, the points

¹⁴https://gitlab.dune-project.org/duneuro/duneuro/-/blob/master/duneuro/test/test_analytical_expressions_local_subtraction.cc?ref_type=heads

¹⁵https://gitlab.dune-project.org/duneuro/duneuro/-/blob/master/duneuro/eeg/local_subtraction_cg_pi_local_operator.hh?ref_type=heads

$\Phi_{A^\top}(p_0), \Phi_{A^\top}(p_1), \Phi_{A^\top}(p_2)$, and $\Phi_{A^\top}(p_3)$ are furthermore affinely independent, and hence $\Phi_{A^\top}(T)$ is again a tetrahedron.

In particular, we thus have that $\Phi_{A^\top}(F)$ is a triangular facet of $\Phi_{A^\top}(T)$ with the corners $\Phi_{A^\top}(p_0), \Phi_{A^\top}(p_1)$, and $\Phi_{A^\top}(p_2)$. Let $\eta_{\Phi_{A^\top}(F)}$ denote its unit outer normal vector. Then, $\eta_{\Phi_{A^\top}(F)}$ has to be orthogonal to $A^\top(p_1 - p_0)$ and $A^\top(p_2 - p_0)$. This implies that $A \cdot \eta_{\Phi_{A^\top}(F)}$ is orthogonal to $p_1 - p_0$ and $p_2 - p_0$. Since the orthogonal complement of $\{p_1 - p_0, p_2 - p_0\}$ in \mathbb{R}^3 is one-dimensional, it thus follows that $\eta_{\Phi_{A^\top}(F)}$ is either $\frac{A^{-1} \cdot \eta}{\|A^{-1} \cdot \eta\|}$ or $-\frac{A^{-1} \cdot \eta}{\|A^{-1} \cdot \eta\|}$. Now, note that $\Phi_{A^\top}(p_3) - \Phi_{A^\top}(p_0)$ points in the opposite direction of $\eta_{\Phi_{A^\top}(F)}$, and $p_3 - p_0$ points in the opposite direction of η . This yields

$$\left\langle \frac{A^{-1} \cdot \eta}{\|A^{-1} \cdot \eta\|}, \Phi_{A^\top}(p_3) - \Phi_{A^\top}(p_0) \right\rangle = \frac{1}{\|A^{-1} \cdot \eta\|} \langle \eta, p_3 - p_0 \rangle < 0.$$

Hence, we have $\eta_{\Phi_{A^\top}(F)} = \frac{A^{-1} \cdot \eta}{\|A^{-1} \cdot \eta\|}$. □

Before computing the analytical expressions, we need to investigate how the integral over a triangular facet transforms under Φ_{A^\top} .

Theorem 2.6.22. *Let $F \subset \mathbb{R}^3$ denote a triangle. Then, the affine hull of the triangle F can be written as $\text{aff}(F) = w + E$, where $E \subset \mathbb{R}^2$ is a two-dimensional linear subspace of \mathbb{R}^3 , and w is an arbitrary point in the affine hull of F . Let p_E denote the orthogonal projection of \mathbb{R}^3 onto E . Then, we define a linear map*

$$(\sigma^\infty)_E^{-1} : E \rightarrow E; y \mapsto p_E((\sigma^\infty)^{-1} \cdot y).$$

Furthermore, let $g : \Phi_{A^\top}(F) \rightarrow \mathbb{R}$ be an integrable function. Then, $g \circ \Phi_{A^\top}$ is integrable over F , and we have

$$\int_{\Phi_{A^\top}(F)} g \, dS = \sqrt{\det((\sigma^\infty)_E^{-1})} \cdot \int_F g \circ \Phi_{A^\top} \, dS.$$

If $\hat{F} \subset \mathbb{R}^2$ is a reference triangle and $\mu : \hat{F} \rightarrow F$ is an affine map of F , we can compute the factor in front of the right-hand side integral via

$$\det((\sigma^\infty)_E^{-1}) = \frac{\det(J_\mu^\top \cdot (\sigma^\infty)^{-1} \cdot J_\mu)}{\det(J_\mu^\top J_\mu)}.$$

Proof. Let $\mu : \hat{F} \rightarrow F$ be an affine map of the triangle F . Then $\Phi_{A^\top} \circ \mu : \hat{F} \rightarrow \Phi_{A^\top}(F)$ is a map of $\Phi_{A^\top}(F)$. If now λ_2 denotes the Lebesgue measure on \mathbb{R}^2 , we have by the definition of the integral over a manifold that

$$\begin{aligned} \int_{\Phi_{A^\top}(F)} g \, dS &= \int_{\hat{F}} g \circ (\Phi_{A^\top} \circ \mu) \cdot \sqrt{\det(J_{\Phi_{A^\top} \circ \mu}^\top J_{\Phi_{A^\top} \circ \mu})} \, d\lambda_2 \\ &= \sqrt{\frac{\det(J_{\Phi_{A^\top} \circ \mu}^\top J_{\Phi_{A^\top} \circ \mu})}{\det(J_\mu^\top J_\mu)}} \cdot \int_{\hat{F}} (g \circ \Phi_{A^\top}) \circ \mu \cdot \sqrt{\det(J_\mu^\top J_\mu)} \, d\lambda_2 \\ &= \sqrt{\frac{\det(J_{\Phi_{A^\top} \circ \mu}^\top J_{\Phi_{A^\top} \circ \mu})}{\det(J_\mu^\top J_\mu)}} \cdot \int_F g \circ \Phi_{A^\top} \, dS. \end{aligned}$$

Now, we have

$$J_{\Phi_{A^\top} \circ \mu}^\top J_{\Phi_{A^\top} \circ \mu} = J_\mu^\top \cdot A \cdot A^\top \cdot J_\mu = J_\mu^\top \cdot (\sigma^\infty)^{-1} \cdot J_\mu.$$

Now, since we assumed μ to be affine, it can be represented as $\mu(z) = M \cdot z + o$, for $M \in \mathbb{R}^{3 \times 2}$ and $o \in \mathbb{R}^3$. Since μ is a map of F , it follows that the columns of M must span the plane E . Furthermore, note that $J_\mu = M$. If now $M = USV^\top$ is a thin singular value decomposition of M , with $U \in \mathbb{R}^{3 \times 2}$, $S \in \mathbb{R}^{2 \times 2}$, and $V \in \mathbb{R}^{2 \times 2}$, it follows that $J_\mu^\top J_\mu = VS^2V^\top$, and thus

$$\begin{aligned} \frac{\det(J_{\Phi_{A^\top \circ \mu}}^\top J_{\Phi_{A^\top \circ \mu}})}{\det(J_\mu^\top J_\mu)} &= \det\left(J_\mu^\top \cdot (\sigma^\infty)^{-1} \cdot J_\mu \cdot (J_\mu^\top J_\mu)^{-1}\right) \\ &= \det\left(VSU^\top (\sigma^\infty)^{-1} US^{-1}V^\top\right) \\ &= \det\left(U^\top (\sigma^\infty)^{-1} U\right). \end{aligned}$$

If we now write $U = (u_1, u_2)$, we have

$$U^\top (\sigma^\infty)^{-1} U = \begin{pmatrix} \langle u_1, (\sigma^\infty)^{-1} u_1 \rangle & \langle u_1, (\sigma^\infty)^{-1} u_2 \rangle \\ \langle u_2, (\sigma^\infty)^{-1} u_1 \rangle & \langle u_2, (\sigma^\infty)^{-1} u_2 \rangle \end{pmatrix}.$$

Since the range of U is the same as the range of M , we have that $\{u_1, u_2\}$ is an orthonormal basis of the plane E . We thus see that $U^\top (\sigma^\infty)^{-1} U$ is the coordinate representation of the linear map $(\sigma^\infty)_E^{-1}$ in the basis $\{u_1, u_2\}$, and hence

$$\det\left(U^\top (\sigma^\infty)^{-1} U\right) = \det\left((\sigma^\infty)_E^{-1}\right).$$

This finishes the proof of the theorem. Note in particular that the value of $\frac{\det(J_{\Phi_{A^\top \circ \mu}}^\top J_{\Phi_{A^\top \circ \mu}})}{\det(J_\mu^\top J_\mu)}$ is independent of the choice of affine map μ . \square

For the following computations, we define

$$c_{\sigma^\infty, F} := \sqrt{\det\left((\sigma^\infty)_E^{-1}\right)}. \quad (2.71)$$

Note that this is well defined, as E is uniquely determined from F .

We now have all the necessary tools to analytically compute the anisotropic local subtraction right-hand side integrals.

Reducing the anisotropic case to the isotropic case

In the following, we aim to present the analytical expressions in a form that lends itself to an easy implementation into DUNEuro, our toolbox for FEM-based EEG and MEG forward simulation [186].

Computing the patch and transition integrals We start with the patch integrals. Let T denote a tetrahedron. In the following, let φ_i , $1 \leq i \leq 4$, denote the unique affine linear map that is 1 on the i -th vertex and 0 on the remaining vertices. Furthermore, let \hat{T} denote the reference tetrahedron, $\hat{\varphi}_i$ the corresponding shape function on the reference tetrahedron, and $\mu_T : \hat{T} \rightarrow T$ an affine map from the reference tetrahedron to T . Then, $\hat{\varphi} = \varphi \circ \mu_T$, and we have by lemma 2.6.3 that $\nabla \varphi_i = J_{\mu_T}^{-\top} \cdot \nabla \hat{\varphi}_i$. Note that, since φ_i is linear, its gradient is a constant function, and we denote this constant vector by $\nabla \varphi_i$. Furthermore, let

$$G := (\nabla \hat{\varphi}_1, \nabla \hat{\varphi}_2, \nabla \hat{\varphi}_3, \nabla \hat{\varphi}_4) \in \mathbb{R}^{3 \times 4}.$$

Remembering that

$$u^\infty = \det(A) \cdot f_{A^\top x_0, A^\top M} \circ \Phi_{A^\top},$$

we can now compute

$$\begin{aligned} \begin{pmatrix} \int_T \langle \sigma^c \nabla u^\infty, \nabla \varphi_1 \rangle dV \\ \int_T \langle \sigma^c \nabla u^\infty, \nabla \varphi_2 \rangle dV \\ \int_T \langle \sigma^c \nabla u^\infty, \nabla \varphi_3 \rangle dV \\ \int_T \langle \sigma^c \nabla u^\infty, \nabla \varphi_4 \rangle dV \end{pmatrix} &= \begin{pmatrix} \nabla \varphi_1^\top \\ \nabla \varphi_2^\top \\ \nabla \varphi_3^\top \\ \nabla \varphi_4^\top \end{pmatrix} \cdot \sigma^c \int_T \nabla u^\infty dV \\ &= (J_{\mu_T}^{-\top} \cdot G)^\top \cdot \sigma^c \cdot A \cdot \int_T \det(A^\top) (\nabla f_{A^\top x_0, A^\top M}) \circ \Phi_{A^\top} dV \\ &= (A^\top \cdot \sigma^c \cdot J_{\mu_T}^{-\top} \cdot G)^\top \cdot \int_{\Phi_{A^\top}(T)} \nabla f_{A^\top x_0, A^\top M} dV \end{aligned}$$

Now, for $1 \leq i \leq 4$ let F_i denote the i -th face of T , and let η_i denote its unit outer normal. By the divergence theorem, we then have

$$\begin{pmatrix} \int_T \langle \sigma^c \nabla u^\infty, \nabla \varphi_1 \rangle dV \\ \int_T \langle \sigma^c \nabla u^\infty, \nabla \varphi_2 \rangle dV \\ \int_T \langle \sigma^c \nabla u^\infty, \nabla \varphi_3 \rangle dV \\ \int_T \langle \sigma^c \nabla u^\infty, \nabla \varphi_4 \rangle dV \end{pmatrix} = (A^\top \cdot \sigma^c \cdot J_{\mu_T}^{-\top} \cdot G)^\top \cdot \sum_{i=1}^4 \left(\int_{\Phi_{A^\top}(F_i)} f_{A^\top x_0, A^\top M} dS \right) \cdot \frac{A^{-1} \cdot \eta_i}{\|A^{-1} \cdot \eta_i\|}.$$

But at this point, we are done, as integrals of the form

$$\int_{\Phi_{A^\top}(F_i)} f_{A^\top x_0, A^\top M} dS$$

were already computed for the isotropic analytical expressions.

We now continue with the transition integrals. First, note that the transformed cutoff function $\chi_{A^{-\top}} := \chi \circ \Phi_{A^{-\top}} : \Phi_{A^\top}(T) \rightarrow \mathbb{R}$ is as a composition of affine maps again affine. We can now, similarly to the patch integral case, compute

$$\begin{aligned} \begin{pmatrix} \int_T \langle \sigma \nabla \chi \cdot u^\infty, \nabla \varphi_1 \rangle dV \\ \int_T \langle \sigma \nabla \chi \cdot u^\infty, \nabla \varphi_2 \rangle dV \\ \int_T \langle \sigma \nabla \chi \cdot u^\infty, \nabla \varphi_3 \rangle dV \\ \int_T \langle \sigma \nabla \chi \cdot u^\infty, \nabla \varphi_4 \rangle dV \end{pmatrix} &= (\sigma \cdot J_{\mu_T}^{-\top} \cdot G)^\top \cdot \int_T \nabla (\chi \cdot u^\infty) dV \\ &= (A^\top \cdot \sigma \cdot J_{\mu_T}^{-\top} \cdot G)^\top \cdot \int_{\Phi_{A^\top}(T)} \nabla (\chi_{A^{-\top}} \cdot f_{A^\top x_0, A^\top M}) dV \\ &= (A^\top \cdot \sigma \cdot J_{\mu_T}^{-\top} \cdot G)^\top \cdot \sum_{i=1}^4 \left(\int_{\Phi_{A^\top}(F_i)} \chi_{A^{-\top}} \cdot f_{A^\top x_0, A^\top M} dS \right) \cdot \frac{A^{-1} \cdot \eta_i}{\|A^{-1} \cdot \eta_i\|}. \end{aligned}$$

But here we are, again, done, as integrals of the form

$$\int_{\Phi_{A^\top}(F_i)} \chi_{A^{-\top}} \cdot f_{A^\top x_0, A^\top M} dS$$

are also already covered by the isotropic case.

Computing the surface integrals Let $F \subset \mathbb{R}^3$ denote a triangle. Let η denote a unit normal vector of F . Let φ_i , $1 \leq i \leq 3$, denote the unique affine linear map that is 1 on the i -th vertex and 0 on the remaining vertices. We start with the surface integral. By theorem 2.6.21, we know that $\eta_{\Phi_{A^\top}(F)} := \frac{A^{-1} \cdot \eta}{\|A^{-1} \cdot \eta\|}$ is a unit normal vector for the triangle $\Phi_{A^\top}(F)$. Let $1 \leq i \leq 3$ be arbitrary. We then have

$$\begin{aligned} \int_F \langle \sigma^\infty \nabla u^\infty, \eta \rangle \varphi_i dS &= \eta^\top \sigma^\infty \int_F (\nabla u^\infty) \cdot \varphi_i dS = \eta^\top \sigma^\infty \det(A) \int_F \nabla (f_{A^\top x_0, A^\top M} \circ \Phi_{A^\top}) \cdot \varphi_i dS \\ &= \eta^\top \sigma^\infty \det(A) A \int_F (\nabla f_{A^\top x_0, A^\top M}) \circ \Phi_{A^\top} \cdot (\varphi_i \circ \Phi_{A^{-\top}}) \circ \Phi_{A^\top} dS \\ &= \frac{\eta^\top \sigma^\infty \det(A) A}{c_{\sigma^\infty, F}} \int_{\Phi_{A^\top}(F)} \nabla f_{A^\top x_0, A^\top M} \cdot (\varphi_i \circ \Phi_{A^{-\top}}) dS. \end{aligned}$$

Now, by definition $(\sigma^\infty)^{-1} = A \cdot A^\top$, and hence

$$\eta^\top \sigma^\infty A = (A^{-1} \eta)^\top = \|A^{-1} \cdot \eta\| \cdot \eta_{\Phi_{A^\top}(F)}^\top.$$

We thus have

$$\int_F \langle \sigma^\infty \nabla u^\infty, \eta \rangle \varphi_i dS = \frac{\det(A) \cdot \|A^{-1} \cdot \eta\|}{c_{\sigma^\infty, F}} \cdot \int_{\Phi_{A^\top}(F)} \langle \nabla f_{A^\top x_0, A^\top M}, \eta_{\Phi_{A^\top}(F)} \rangle \cdot (\varphi_i \circ \Phi_{A^{-\top}}) dS.$$

But now notice that $\varphi_i \circ \Phi_{A^{-\top}}$ is the unique affine linear function that is 1 on the i -th corner of $\Phi_{A^\top}(F)$ and 0 on the remaining corners of $\Phi_{A^\top}(F)$. In particular, we thus see that

$$\int_{\Phi_{A^\top}(F)} \langle \nabla f_{A^\top x_0, A^\top M}, \eta_{\Phi_{A^\top}(F)} \rangle \cdot (\varphi_i \circ \Phi_{A^{-\top}}) dS$$

is exactly the surface integral I_S for unit conductivity on the triangle $\Phi_{A^\top}(F)$ with the affine trial function $\varphi_i \circ \Phi_{A^{-\top}}$, and we have thus reduced the computation of the anisotropic surface integral to the computation of an isotropic surface integral.

For the electrode interface integral, we can similarly compute

$$\begin{aligned} \int_F \chi \cdot u^\infty \cdot \varphi_i dS &= \det(A) \cdot \int_F \chi_{A^{-\top}} \circ \Phi_{A^\top} \cdot f_{A^\top x_0, A^\top M} \circ \Phi_{A^\top} \cdot (\varphi_i \circ \Phi_{A^{-\top}}) \circ \Phi_{A^\top} dS \\ &= \frac{\det(A)}{c_{\sigma^\infty, F}} \cdot \int_{\Phi_{A^\top}(F)} \chi_{A^{-\top}} \cdot f_{A^\top x_0, A^\top M} \cdot (\varphi_i \circ \Phi_{A^{-\top}}) dS. \end{aligned}$$

We see that this last integral is exactly the integral that arises in the isotropic case with unit conductivity for the electrode interface integrals. We have thus reduced the anisotropic electrode interface integrals to the isotropic case.

By almost the same computation, we have

$$\int_F \chi \cdot u^\infty dS = \frac{\det(A)}{c_{\sigma^\infty, F}} \cdot \int_{\Phi_{A^\top}(F)} \chi_{A^{-\top}} \cdot f_{A^\top x_0, A^\top M} dS,$$

which reduces the electrode DOF integral to the isotropic case.

Implementation All formulas derived in this subsection have been implemented into the DUNEuro toolbox. For a test validating the formulas against numerical quadrature see here,¹⁶ and for the implementation into a DUNE local operator see here.¹⁷ In this implementation, we compute the decomposition $(\sigma^\infty)^{-1} = A \cdot A^\top$ by computing a Cholesky decomposition.

2.7 Further Implementation Aspects

The derivation of the analytical expressions for the local subtraction right-hand side, as shown in the preceding section, is the most involved part in the implementation of the local subtraction approach. There are, however, four remaining points we want to discuss before going to numerical evaluations.

- For the local subtraction approach, we need to construct a patch Ω^∞ , a transition region $\tilde{\Omega}$, and a cutoff function χ . In section 2.7.1, we want to present a strategy for how to perform this construction.
- For the local subtraction MEG forward problem, we do not have access to analytical formulas. We thus need to employ numerical integration. The natural question is then, of course, what integration order we should choose. This is investigated in section 2.7.2.
- If we want to reconstruct the neural activity underlying a measurement, we need a set of candidate source positions. It is well known that the choice of source positions can have a significant impact on the reconstruction [216]. If one, e.g., uses the Venant approach for forward modeling, one should ensure that each source position fulfills the so-called *Venant condition*, meaning that the closest mesh vertex is exclusively contained in gray matter elements [130]. But while this might be an appropriate condition for the Venant approach, it still allows degenerate source positions that are extremely close to a conductivity jump. We instead suggest ensuring that the candidate positions have a sufficient distance from the brain compartment boundary. We hence want to discuss how to implement the construction of such a source space, which is the topic of section 2.7.3.
- We use DUNEuro for all our numerical computations, which is a C++ toolbox based on the DUNE framework [186]. Development on this toolbox started in 2016, and the corresponding GitLab repository¹⁸ contains, at the point of writing, commits from 17 different contributors. The DUNEuro toolbox is thus a collaborative effort, encompassing a large set of features with occasionally complex interdependencies, not all of which are understood by each contributor. Given this it is of utmost importance to set up an environment capable of automatically compiling and testing the code, to ensure that a new change does not break existing functionality, and to guarantee that the code does not only work on the programmers machine with the programmers compiler setup of choice, but also on a larger set of targeted machines and toolchains. Somewhat surprisingly, no such environment existed for DUNEuro up until the fall of 2024. We spend considerable effort on setting up a so-called *continuous integration (CI) pipeline*, to enable the automatic running of tests on different environments. In section 2.7.4, we want to briefly present the outcome of this work.

We will now discuss these points.

¹⁶https://gitlab.dune-project.org/duneuro/duneuro/-/blob/master/duneuro/test/test_anisotropic_analytical_expressions_local_subtraction.cc?ref_type=heads

¹⁷https://gitlab.dune-project.org/duneuro/duneuro/-/blob/master/duneuro/eeg/local_subtraction_cg_pi_anisotropic_local_operator.hh?ref_type=heads

¹⁸<https://gitlab.dune-project.org/duneuro>

2.7.1 Constructing the local subtraction approach

The local subtraction approach, as defined in definition 2.5.12, is based on a choice of patch Ω^∞ , transition region $\tilde{\Omega}$, and cutoff function χ . We propose to construct these objects in the same way as we suggested in [91]. Concretely, since we want to apply a FEM approach, we can assume that we have access to a mesh. We then first locate the element containing the source position. Note that this can be done efficiently using *edgehopping*, as described in [186]. The element containing the source is our starting point for constructing the patch. In the end, both the patch and the transition region will consist of a union of mesh elements.

We then build the patch by performing so-called *vertex extensions*. If we have a set \mathcal{T} of mesh elements, the vertex extension $V_E(\mathcal{T})$ of this set is given by the set of mesh elements that share at least one vertex with an element from \mathcal{T} . Note that, since each element shares a vertex with itself, we have $\mathcal{T} \subset V_E(\mathcal{T})$. We then choose Ω^∞ by repeatedly applying vertex extensions, i.e. if e is the element containing the dipole, and $\mathcal{T}_0 = \{e\}$, we set

$$\Omega^\infty := \bigcup V_E^n(\mathcal{T}_0),$$

where $V_E^n(\mathcal{T}_0)$ means V_E is applied n times to \mathcal{T}_0 . Then, we set

$$\tilde{\Omega} := \bigcup (V_E^{n+1}(\mathcal{T}_0) \setminus V_E^n(\mathcal{T}_0)),$$

and, in the case of tetrahedral meshes, we choose χ to be the unique element-wise affine function that is 1 on every patch vertex and 0 on the remaining mesh vertices. The function χ then has the properties that $\chi = 1$ on Ω^∞ , $\chi = 0$ on $\Omega \setminus (\tilde{\Omega} \cup \Omega^\infty)$, and that $\chi \in H^1(\Omega)$. Furthermore, as long as $n \geq 1$, it follows that Ω^∞ is an environment of the source position. This construction of Ω^∞ , $\tilde{\Omega}$, and χ thus fulfills the requirements of the local subtraction approach.

One question that still needs to be addressed is the choice of n . From an efficiency point of view, one wants to choose n as small as possible, as larger n 's lead to larger patches and transition regions, which lead to a more involved computation. On the other hand, note that on the patch Ω^∞ we solve for $u^c = u - u^\infty$, while outside of $\tilde{\Omega} \cup \Omega^\infty$ we solve for $u^c = u$. Close to the source position, the potential u shows a singular behaviour, and if we choose Ω^∞ too small, we force some of the singular behaviour of u onto u^c instead of subtracting it out, which might be disadvantageous for the FEM simulation. This suggests that the patch should not be chosen too small. In contrast, however, choosing the patch to be too large can also degrade the accuracy of the approach, as the infinity potential u^∞ completely ignores the complex conductivity distribution in the head, which significantly distorts the true potential u when compared to the homogeneous potential u^∞ . This can lead to a complex structure in $u - \chi \cdot u^\infty$ that is difficult to properly resolve in a FEM simulation. Imagine, e.g., the brain-CSF interface, where the highly conducting CSF has a strong channeling effect on the current density represented by $-\sigma \nabla u$, or the fact that the insulating effect of the skull causes the potential u to be much weaker in the skin than the homogeneous potential u^∞ .

In section 2.8.2, we will use numerical simulations to investigate the impact of these effects. The result will be that choosing $n = 2$ seems to be the optimal choice for the meshes that we typically employ in practice. Note that the visualization in figure 2.6 also, not coincidentally, corresponds to performing two vertex extensions to construct the patch.

2.7.2 Integration orders for MEG postprocessing

The content of this subsection is largely the same as the corresponding investigation in the supplementary material of our publication [91]. In fact, we will focus on tetrahedral meshes in this thesis. If the reader is interested in integration orders for hexahedral meshes, we refer to [91].

As summarized in definition 2.5.17, to solve the MEG forward problem, we need, for tetrahedral elements K and triangles $F \subset \mathbb{R}^3$, to assemble the following integrals

$$F_P(x) = \int_K \sigma^c(y) \nabla u^\infty(y) \times \frac{x-y}{\|x-y\|^3} dV(y), \quad (2.72)$$

$$F_B(x) = \int_F \sigma^\infty(\chi \cdot u^\infty(y)) \cdot \eta(y) \times \frac{x-y}{\|x-y\|^3} dS(y), \quad (2.73)$$

$$F_T(x) = \int_K \sigma^c(y) \nabla (\chi \cdot u^\infty)(y) \times \frac{x-y}{\|x-y\|^3} dV(y), \quad (2.74)$$

where we call F_P *patch flux*, F_B *boundary flux*, and F_T *transition flux*. Here, u^∞ is again the solution of the dipole problem in an unbounded homogeneous conductor for a dipole at position x_0 with moment M , as derived in (2.56). We again emphasize that, for the MEG forward problem, we always assume an isotropic conductivity in the source element, while non-source elements can still be anisotropic.

Our heuristic for choosing integration orders is inspired by [18], where the authors empirically shows that the relative error introduced by numerical integration only depends on the ratio $\frac{d}{a}$, where d denotes the distance of the dipole position x_0 from the entity we are integrating over, and a denotes the maximum edge length of the entity. For a fixed integration order, one generally has that a larger ratio corresponds to a smaller error. More concretely, if either the source gets closer to the integration entity or the integration entity gets larger, the accuracy of the numerical integration typically degrades.

One can, in fact, also give a mathematical argument for why only the ratio of these two values is important for the accuracy of the numerical integration. To perform numerical integration, one typically chooses a so-called *quadrature rule* on some reference element, say \hat{T} . A quadrature rule on \hat{T} is given by a set $\{(z_1, w_1), \dots, (z_K, w_K)\}$, where $z_i \in \hat{T}$ and $w_i \in \mathbb{R}$, and for a function $f : \hat{T} \rightarrow \mathbb{R}$ one approximates

$$\int_{\hat{T}} f dV \approx \sum_{i=1}^K w_i \cdot f(z_i).$$

Quadrature rules are typically characterized by their *order*, which is defined to be the largest $d \in \mathbb{N}_0$ such that the “ \approx ” in the sum above is in fact a “=” for all polynomials of degree less than or equal to d . If we now have a function f on the tetrahedron T , we approximate its integral by choosing an affine map μ_T of T from the reference element, and computing

$$\int_T f dV = \int_{\hat{T}} f \circ \mu_T \cdot |\det(J_{\mu_T})| dV \approx \sum_{i=1}^K w_i \cdot |\det(J_{\mu_T})| \cdot f(\mu_T(z_i)). \quad (2.75)$$

We now want to investigate the effect of scaling on the error introduced by numerical integration. To this end, let $0 \neq \lambda \in \mathbb{R}$. Then, scaling the setup by λ leads to a source at position $\lambda \cdot x_0$ and a coil position at $\lambda \cdot x$, which we integrate over the element $\lambda \cdot K$. But then, $\lambda \cdot \mu_T$ is a map of $\lambda \cdot K$, and we have $u_{\lambda x_0}^\infty(\lambda z) = \frac{1}{\lambda^2} \cdot u_{x_0}^\infty(z)$, $\nabla u_{\lambda x_0}^\infty(\lambda z) = \frac{1}{\lambda^3} \cdot \nabla u_{x_0}^\infty(z)$, and $\det(J_{\lambda \mu_T}) = \lambda^3 \det(J_{\mu_T})$.

Furthermore, in practice, we always choose χ to be linear on each element. Inserting these observations into (2.75) for f either the integrand of the patch flux (2.72) or the transition flux (2.74), one sees that both the integral as well as the numerical quadrature result are scaled by a factor of $\frac{1}{\chi^2}$ when scaling the setup by a factor of λ . This, in particular, implies that the relative error between the numerical quadrature and the true value of the integrals does not change when a scaling is applied. A similar argument shows that the same is true for the boundary flux in (2.73). Finally, note that scaling either the dipole moment or the conductivity also does not change the relative error introduced by numerical integration.

But since scaling does not affect the relative accuracy of numerical integration, we can simply investigate the integration error at some reference scale and then, in practice, choose a suitable quadrature rule by comparing with the results at the reference scale. To this end, we implemented a test where we took a tetrahedron of maximal side length one and investigated the accuracy of the numerical quadrature for source positions at different distances to the tetrahedron.¹⁹ Our tests suggests that, when compared to the distance of the position to the tetrahedron, the concrete position of the dipole, the shape of the tetrahedron, and the dipole moment had a comparatively small impact on the accuracy of the numerical integration. We thus arrive at the same conclusion as [18], i.e., that the ratio $\frac{d}{a}$ should be used to select the order of the quadrature rule.

The conclusion of section 2.7.1 was to perform two vertex extensions to construct the patch. Thus, one can expect that there are two element layers between the source positions and the transition elements, and thus that for the transition flux we have, on average, a ratio $\frac{d}{a}$ of 2. To get a better intuition for the distribution of ratios, we took a multilayer sphere mesh and 1000 dipole positions close to the CSF boundary, and, for each of these positions, computed the ratio $\frac{d}{a}$ for each transition element. The resulting distribution is visualized in figure 2.7.

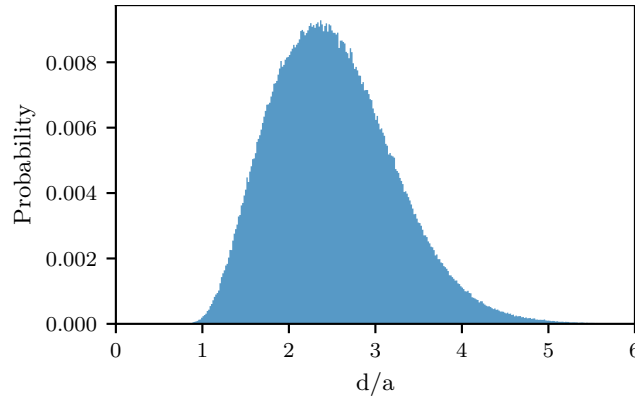


Figure 2.7: Distribution of ratios $\frac{d}{a}$, where d is the distance of the source position from the element and a is the maximal edge length of the element, sampled over the transition elements of 1000 dipolar sources. The patches were constructed by performing two vertex extensions, and the transition region by performing one additional extension. This figure is taken from the supplementary material of [91].

We see that the distribution indeed peaks close to a ratio of 2. Furthermore, we see that almost all elements correspond to a ratio of $\frac{d}{a} \geq 1$. In particular, if we choose the integration order in

¹⁹https://github.com/MalteHoe1/local_integration_test/blob/master/dune/local_integration_test/local_integration_test.hh

such a way that they are accurate up until ratios of slightly below 1, we can expect an accurate assembly of the local subtraction MEG transition fluxes.

Furthermore, our experience suggests that targeting a worst-case relative error of 10^{-3} for the numerical quadrature is sufficient for the MEG forward modeling errors not to be dominated by the integration error. To now gauge the integration order for the transition fluxes, we took a tetrahedron with the corners

$$\begin{aligned} p_1 &= \left(0, -\frac{1}{2}, -\frac{\sqrt{3}}{6}\right), \\ p_2 &= \left(0, \frac{1}{2}, -\frac{\sqrt{3}}{6}\right), \\ p_3 &= \left(0, 0, -\frac{\sqrt{3}}{3}\right), \\ p_4 &= \left(-\frac{\sqrt{6}}{3}, 0, 0\right), \end{aligned}$$

placed a source position at $x_0 = d \cdot (1, 0.1, 0.1)$ for different values of d and looked at a coil position at $x = (20, 0, 0)$. Note that the tetrahedron with the corners p_1, p_2, p_3 , and p_4 has equal side lengths of 1. Furthermore, we explicitly did not place the source position on the x axis, as in this case some integration orders, at some distances, perform extraordinarily well, similar to what is shown in figure 2a) in [18]. We suspect that this is due to $(1, 0, 0)$ lying directly above the barycenter of the face with the vertices p_1, p_2 , and p_3 of the tetrahedron, which produces a highly symmetric scenario from which some integration orders might benefit.

We used the quadrature rules implemented in the DUNE geometry module. For tetrahedrons, optimized rules up to an order of 5 are implemented, while, for orders larger than 5, quadrature rules are constructed from products of 1-dimensional rules. We refer to [176] and [198] for details. The resulting error curves are shown in figure 2.8. Based on this figure, we suggest using an integration order of 5 for the MEG transition fluxes.

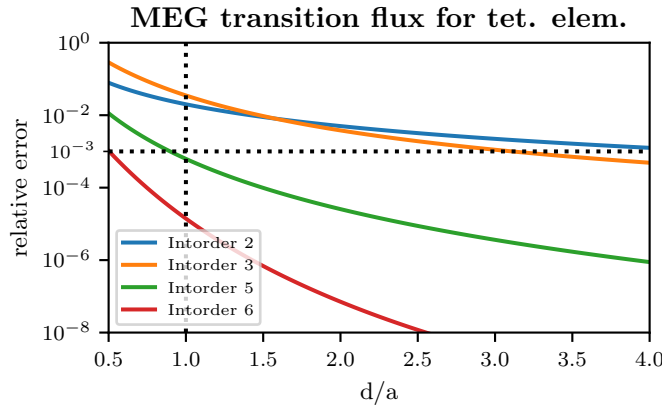


Figure 2.8: Relative errors for the numerical integration of the MEG transition flux for different integration orders and different distance-to-edge-length ratios $\frac{d}{a}$. Errors were computed by comparing against an integration order of 50. Note that we did not include a rule of order 4, as requesting a rule of order 4 from the DUNE geometry module yields the same result as requesting an order 5 rule. This figure is taken from the supplementary material of [91].

For the MEG patch flux, the situation is more complicated. The problem is that we cannot

Table 2.1: Tetrahedral MEG patch flux integration order

$\frac{\text{distance}}{\text{edge length}} \left(\frac{d}{a}\right)$	$\frac{d}{a} \geq 0.5$	$\frac{d}{a} \geq 0.4$	$\frac{d}{a} \geq 0.33$	$\frac{d}{a} \geq 0.25$	$\frac{d}{a} \geq 0.17$
Intorder	8	9	11	13	20

expect the distance d from the source position to the element to be nicely bounded below. We instead suggest computing, for each patch element that needs to be assembled, the distance of the dipole position from the element and adaptively choosing an integration order that is sufficiently accurate for the present distance-edge-length ratio. Based on the results in figure 2.9, we thus suggest choosing the quadrature rule as described by table 2.1. This covers ratios $\frac{d}{a}$ that are greater than $\frac{1}{6}$. If we e.g. assume that the elements in the patch have an edge length of at most 2 mm, this implies that we expect the integration to be accurate as long as the distance of the dipole position from the brain compartment boundary is at least 0.34 mm. Given that the physiologically most relevant neural sources are expected to have a distance between 1 mm and 2 mm to a conductivity jump [152], we argue that it is reasonable to expect that a collection of potential source positions fulfills this condition.

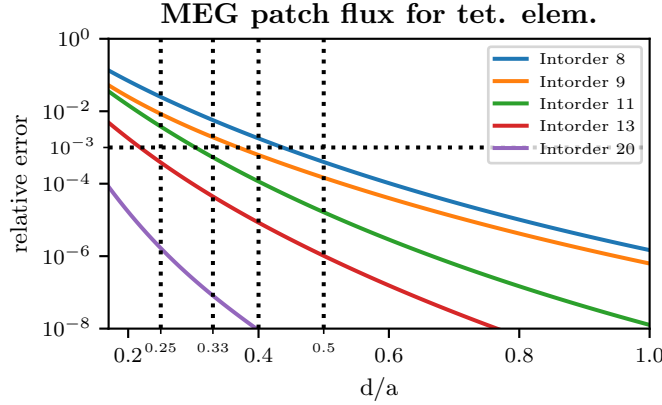


Figure 2.9: Relative errors for the numerical integration of the MEG patch flux for different integration orders and different distance-to-edge-length ratios $\frac{d}{a}$. Errors were computed by comparing against an integration order of 50. This figure is taken from the supplementary material of [91].

Finally, we want to discuss the MEG boundary flux from (2.73). We expect that an assembly of this integral will almost never be necessary, and will hence not show an additional plot here. In [91], it is shown that when evaluating MEG flux integrals on the patch boundary, an integration order of 6 is sufficient when performing two vertex extensions. Since typically the domain boundary will be further away from the source position than the patch boundary arising from two vertex extensions, we argue that an order of 6 is also a sensible default value for the MEG boundary flux. Note, however, that if there is a reason to expect that there is a domain boundary close to the source position, one should overwrite this default, and probably use a strategy similar to the one given in table 2.1.

To summarize, we suggest using an integration order of 5 for the MEG transition flux, an integration order of 6 for the MEG boundary flux, and an integration order based on table 2.1 for

the MEG patch flux. This is also the current default choice implemented in DUNEuro.²⁰

2.7.3 Source space construction

When trying to reconstruct neural activity, many inverse methods, and in particular the ones we typically use, require the construction of a so-called *source space*, i.e., a set of possible locations of neural activity. There are, roughly speaking, two strategies in the community to perform this construction.

1. Place a regular three-dimensional grid over the head model and filter the grid positions by some set of criteria.
2. Estimate a surface going through the cortex and place sources on this surface.

Looking at a realistic head model, e.g. in figure 2.3b, should convince the reader that placing a surface through the middle of the cortex is a delicate task, and while excellent open source software for creating such a surface exists, such as e.g. FreeSurfer [189], it is not always possible to automatically generate an adequate surface. Placing a regular grid, on the other hand, is both easy to implement and, given a sufficient resolution, guaranteed to place sources in all relevant parts of the cortex. In many applications, we try to reconstruct neural activity using a so-called *dipole scan*, where we simply iterate over each source position candidate and check how well a single dipole at this position can explain, in an L2 sense, the measured data. As for this inverse approach at least, we do not see a principal advantage of one source space construction method over the other, we typically prefer strategy 1) over strategy 2) due to its simplicity.

When placing such a regular grid, we need to decide on how to filter out unsuitable source positions. First of all, we of course only allow source positions inside the gray matter compartment. This, however, is typically not sufficient. One central property of numerical EEG and MEG forward modeling is that errors typically get larger the closer the source position is to a conductivity jump, see e.g. [18], [239], or even our own investigations in section 2.8 of this thesis. One thus typically tries to avoid placing sources too close to a conductivity jump. One strategy to achieve this, which was, to the best of our knowledge, first suggested by [130], consists of the so-called *Venant condition*. Here, one checks if the mesh node closest to the candidate source position is exclusively contained in gray matter elements. Candidate source positions that do not fulfill this condition are then rejected. At least in our work group at the IBB, the Venant condition has become the de facto standard, and is used in multiple recent publications [6, 105, 56]. Here, we want to demonstrate that the Venant approach is a rather crude tool to enforce a distance to a conductivity jump that will sometimes completely fail to do so, and that, especially for the local subtraction approach, a different condition should be employed.

First of all, note that the Venant condition can be fulfilled even if the candidate source position is extremely close to a conductivity jump. This can, e.g., happen if the source position lies inside a “flat” tetrahedron that shares one face with an element of a different conductivity and where the interior mesh vertex is close to the conductivity jump. A two-dimensional visualization of this scenario is shown in figure 2.10.

While all FEM forward modeling approaches run into numerical difficulties when the sources get too close to conductivity jumps, this is especially noticeable for the MEG local subtraction approach. The reason for this is that solving the local subtraction MEG forward problem requires the numerical integration of certain integrals involving the unbounded homogeneous medium

²⁰https://gitlab.dune-project.org/duneuro/duneuro/-/blob/master/duneuro/eeg/local_subtraction_source_model.hh?ref_type=heads

Failure of the Venant condition

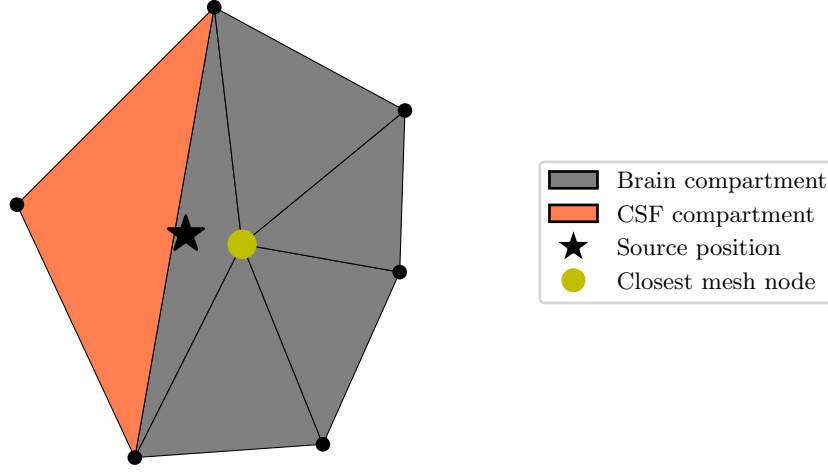


Figure 2.10: 2-dimensional illustration of a scenario where the source position is contained in the brain compartment and the Venant condition is fulfilled, but the source position is still extremely close to a conductivity jump. The gray elements correspond to brain elements, while the orange element corresponds to a CSF element.

solution u^∞ , see definition 2.5.17 and section 2.7.2. Concretely, if K is a patch element that is not contained in the brain compartment, we need to evaluate expressions of the form

$$\int_{\Omega^\infty} \langle \sigma^c \nabla u^\infty \times \frac{x-y}{\|x-y\|^3}, v \rangle dy,$$

where ∇u^∞ , as derived in equation (2.44), behaves like $\frac{1}{\|y-x_0\|^3}$ close to the source position x_0 . The reader can probably imagine that performing numerical integration close to a non-integrable singularity is a numerically difficult task. In section 2.7.2, we derived an integration strategy such that if we build the mesh in such a way that the patch elements can be assumed to have a maximum edge length of 2 mm and the dipole position has a distance of at least 0.34 mm from the conductivity jump, we can expect an accurate evaluation of the integral. But if these conditions are violated, the evaluation of the integral can fail. In fact, if you compute a MEG lead field and see an isolated spike in the lead field magnitude, the distance of the corresponding source position to the conductivity jump is one of the first things you might check.

From our perspective, the natural response to this problem is to directly enforce a minimal distance in the source space construction, instead of hoping that an indirect approach, such as the Venant condition, will naturally lead to enough “good” source positions. The implementation and testing of this idea is the content of the remainder of this subsection. Furthermore, we will again focus on tetrahedral head models.

For tetrahedral head models, the boundary of the brain compartment consists of a set of triangles in \mathbb{R}^3 . Then, the distance of a source position candidate from this compartment boundary can be computed as the minimum of the distances of the position from the triangles making up the

boundary surface. If we have a single triangle, say $F \subset \mathbb{R}^3$, and a source position candidate x_0 , it is relatively straightforward to compute the distance of x_0 from F , e.g., as follows.

- First note that it suffices to find the closest point $q \in F$ to x_0 , as the distance is then simply given by $\|x_0 - q\|_2$.
- Then note that if $E \subset \mathbb{R}^3$ denotes the plane spanned by the triangle F , the closest point in F to x_0 is the same as the closest point in F to the orthogonal projection of x_0 onto E , denoted by $p_E(x_0)$.
- As $p_E(x_0) \in E = \text{aff}(F)$, it can be expressed in affine coordinates as $p_E(x_0) = \sum_{i=1}^3 \lambda_i p_i$, where p_1, p_2 , and p_3 denote the vertices of the triangle and $\sum_{i=1}^3 \lambda_i = 1$. If all λ_i are non-negative, $p_E(x_0)$ is contained in the triangle, and we have found the closest point. If at least one λ_i is negative, $p_E(x_0)$ is outside of the triangle F .
- If $p_E(x_0)$ is outside of F , the closest point in F to $p_E(x_0)$ is contained in an edge of F . Furthermore, one can quickly check that the closest point must in fact be contained in an edge of F that lies opposite to a triangle vertex p_i such that for the associated affine coordinate λ_i of $p_E(x_0)$ we have $\lambda_i < 0$, and we hence only need to further process those edges.
- For an edge, the closest point to $p_E(x_0)$ in this edge is the same as the closest point in the edge to the projection of $p_E(x_0)$ onto the line defined by the edge. But depending on where this projection falls, this is either the projection itself or one of the two endpoints of the edge.

This procedure has been implemented by us into DUNEuro.²¹ The naive algorithm to enforce a distance condition during source space creation is now as follows. First, extract the brain compartment boundary. Then, iterate over each source position candidate, and for each candidate, iterate over each compartment boundary triangle and compute the distance using the procedure outlined above, and get the distance of the source position from the compartment boundary by taking the minimum over all the triangle distances. This does indeed work. The problem, however, lies in the computational costs of this approach. For the realistic head model from figure 2.3b, e.g., the brain compartment boundary consists of ca. 500k triangles. Furthermore, if we place a regular grid with a 2 mm resolution, as e.g. [56] did, we typically get (way) more than 10k source positions. Using the naive approach to filter out source positions based on their distance to the brain compartment boundary then takes multiple hours on contemporary hardware.

To overcome this, we implemented a so-called *bounding volume hierarchy* into DUNEuro.²² Assume that we have a set of geometrical entities $\mathcal{T} = \{e_1, \dots, e_N\}$, where we assume that for $1 \leq i \leq N$ we have $e_i \subset \mathbb{R}^3$.²³ We now want to build a tree structure to describe the geometry of \mathcal{T} . A node, say \mathcal{N} , in this tree is then given by

1. A subset $\mathcal{T}_{\mathcal{N}} \subset \mathcal{T}$ of geometrical entities.
2. A volume $\Omega_{\mathcal{N}} \subset \mathbb{R}^3$ with the property that for each $e \in \mathcal{T}_{\mathcal{N}}$ we have $e \subset \Omega_{\mathcal{N}}$. Such an $\Omega_{\mathcal{N}}$ is then called a *bounding volume*.

²¹https://gitlab.dune-project.org/duneuro/duneuro/-/blob/master/duneuro/common/distance_utilities.hh?ref_type=heads

²²https://gitlab.dune-project.org/duneuro/duneuro/-/blob/master/duneuro/common/bounding_volume_hierarchy.hh?ref_type=heads

²³Note that for the abstract discussion, we could also have chosen \mathbb{R}^n for some $n \in \mathbb{N}$. We will, however, only use the case $n = 3$ in this thesis.

3. If the node is not a leaf node, two references to child nodes $\mathcal{N}_{\text{left}}$ and $\mathcal{N}_{\text{right}}$.

Bounding volume hierarchies are well-known objects in graphics programming, and can be used to accelerate the computation of numerous tasks, among them, in particular, distance computations. Bounding volume hierarchies, and various aspects of their implementation, are an area of active research [242].²⁴ For our purposes, however, the most basic variant is already sufficient.

Concretely, we will always choose $\Omega_{\mathcal{N}}$ to be the axis-aligned bounding box of the entities in $\mathcal{T}_{\mathcal{N}}$. More formally, we set for $1 \leq i \leq 3$

$$a_i^{\mathcal{N}} := \inf_{e \in \mathcal{T}_{\mathcal{N}}} \inf_{x \in e} x_i$$

$$b_i^{\mathcal{N}} := \sup_{e \in \mathcal{T}_{\mathcal{N}}} \sup_{x \in e} x_i$$

and then set the bounding volume as the Cartesian product

$$\Omega_{\mathcal{N}} := [a_1^{\mathcal{N}}, b_1^{\mathcal{N}}] \times [a_2^{\mathcal{N}}, b_2^{\mathcal{N}}] \times [a_3^{\mathcal{N}}, b_3^{\mathcal{N}}].$$

For the root node $\mathcal{N}_{\text{root}}$, we simply choose $\mathcal{T}_{\mathcal{N}_{\text{root}}} = \mathcal{T}$. We then decide on some number $k \in \mathcal{N}$, which we call the *maximum leaf size*²⁵, and build up the tree by starting the following recursive procedure on the root node $\mathcal{N}_{\text{root}}$.

1. Let \mathcal{N} denote the current node. If the number of entities in $\mathcal{T}_{\mathcal{N}}$ is at most k , terminate the recursion.
2. If $\mathcal{T}_{\mathcal{N}}$ contains more than k entities, compute the axis i_0 of $\Omega_{\mathcal{N}}$ with the longest extent, i.e.

$$i_0 = \arg \min_{1 \leq i \leq 3} b_i^{\mathcal{N}} - a_i^{\mathcal{N}}.$$

Then, split the set $\mathcal{T}_{\mathcal{N}}$ into two sets $\mathcal{T}_{\mathcal{N}_{\text{left}}}$ and $\mathcal{T}_{\mathcal{N}_{\text{right}}}$ by the following scheme. For each entity $e \in \mathcal{T}_{\mathcal{N}}$, choose a point $x \in e$. Then, if $x_{i_0} \leq \frac{a_{i_0}^{\mathcal{N}} + b_{i_0}^{\mathcal{N}}}{2}$, assign e to $\mathcal{T}_{\mathcal{N}_{\text{left}}}$, and else assign e to $\mathcal{T}_{\mathcal{N}_{\text{right}}}$. In our application, e will always be a triangle, and we choose x to be the barycenter of the triangle.

3. If either $\mathcal{T}_{\mathcal{N}_{\text{left}}}$ or $\mathcal{T}_{\mathcal{N}_{\text{right}}}$ is empty after the previous step, we terminate the recursion. Else, we start at step 1) for the two new nodes $\mathcal{N}_{\text{left}}$ and $\mathcal{N}_{\text{right}}$.

Once this recursion is done, we have constructed a bounding volume hierarchy. The utility of this data structure for our problem now arises as follows. Assume that we have a reasonably close upper bound d for the distance of a source position candidate x_0 from the set of triangles $\mathcal{T} = \{e_1, \dots, e_N\}$. If we now have some node \mathcal{N} , and we can show that

$$d \leq \text{dist}(x_0, \Omega_{\mathcal{N}}),$$

then we have for each $e \in \mathcal{T}_{\mathcal{N}}$ that $e \subset \Omega_{\mathcal{N}}$, which implies

$$d \leq \text{dist}(x_0, \Omega_{\mathcal{N}}) \leq \text{dist}(x_0, e).$$

Thus, our distance estimate cannot be improved by triangles from $\mathcal{T}_{\mathcal{N}}$, and we can skip all these triangles in our search for the closest triangle. This allows us to quickly disregard large subsets

²⁴While not being an academic research article, I also want to highly recommend <https://www.youtube.com/watch?v=C1H4zIiC0aI> for a truly beautiful introduction to bounding volume hierarchies.

²⁵In our implementation, we chose $k = 100$.

of the total set of triangles and will, in the end, massively improve the efficiency of the source space construction.

As a first step towards such an efficient implementation, notice that, as $\Omega_{\mathcal{N}}$ is an axis-aligned bounding box, $\text{dist}(x_0, \Omega_{\mathcal{N}})$ is easy to compute. If $a, b \in \mathbb{R}$ with $a < b$, set

$$\text{clip}_{[a,b]} : \mathbb{R} \rightarrow [a, b]; x \mapsto \begin{cases} a & \text{if } x \leq a \\ x & \text{if } a < x < b \\ b & \text{if } x \geq b \end{cases}$$

It is then an easy exercise to show that, for an arbitrary $x = (x_1, x_2, x_3) \in \mathbb{R}^3$, the closest point in the Cartesian product $[a_1, b_1] \times [a_2, b_2] \times [a_3, b_3]$ to the point x is given by the point

$$\begin{pmatrix} \text{clip}_{[a_1, b_1]}(x_1) \\ \text{clip}_{[a_2, b_2]}(x_2) \\ \text{clip}_{[a_3, b_3]}(x_3) \end{pmatrix} \in \mathbb{R}^3.$$

And while this is a little unwieldy to write down, it is extremely easy to implement, and we see that computing the distances $\text{dist}(x_0, \Omega_{\mathcal{N}})$ is quick and easy.

We can now compute the distance of a candidate source position x_0 from the set of boundary triangles in the following way. Initialize the current best distance $d = \infty$ and start at the root node.

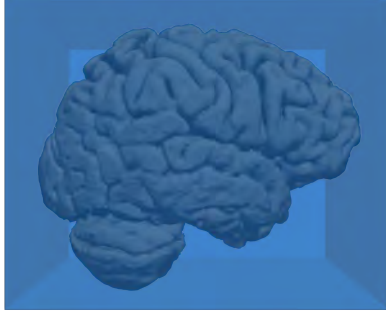
1. If we are in a leaf node, iterate over each triangle in $\mathcal{T}_{\mathcal{N}}$ and compute the distance of the candidate position from the triangle. If a triangle improves on the current best distance d , update the value of d .
2. If we are not in a leaf node, compute the distances of the candidate source position x_0 from the bounding boxes of the child nodes.
 - (a) Start with the child node with the closer bounding box. If the distance from its bounding box is greater than the current best estimate d , we can stop the recursion. Else, we go to step 1) with this child node.
 - (b) Continue with the child node with the more distant bounding box. If the distance from its bounding box is greater than the current best estimate d , we can stop the recursion. Else, we go to step 1) with this child node.

In the strategy above, it is important to search the node with the closer bounding box first. The effect of this is that at the start of the algorithm, we descend down the tree in a greedy, i.e., locally optimal, fashion, and thus quickly get a reasonable upper bound for the distance, allowing us to quickly reject large portions of the bounding volume hierarchy.

In DUNEuro, we implemented a parallelized source space construction using the algorithm outline above.²⁶ On an AMD Ryzen Threadripper 3960X processor with 24 cores, the extraction of the gray matter compartment boundary, the construction of a bounding volume hierarchy on the extracted triangles, and the parallelized checking of containment in the gray matter and the distance condition on a regularly placed grid with a resolution of 1.5 mm took a combined 4.93 s for the realistic mesh from figure 2.3b. We want to emphasize that this does not include the time needed to set up the UG grid manager [13], which took ca. 90 s, and is a task that always

²⁶https://gitlab.dune-project.org/duneuro/duneuro/-/blob/master/duneuro/common/source_space_factory.hh?ref_type=heads

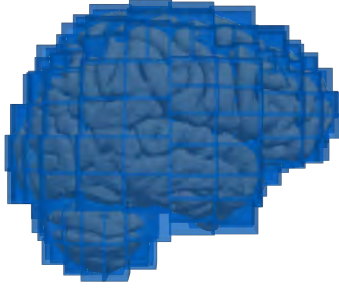
needs to be done when setting up a DUNEuro driver. The resulting bounding volume hierarchy is visualized in figure 2.11.



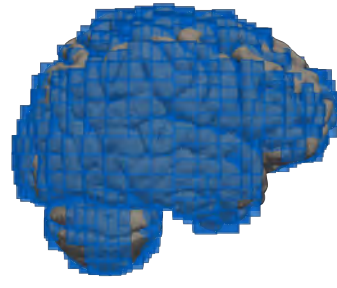
(a) Depth 0



(b) Depth 3



(c) Depth 9



(d) Depth 12

Figure 2.11: Visualization of a bounding volume hierarchy for a set of triangles making up a brain compartment boundary. These triangles were extracted from the realistic volume conductor model shown in figure 2.3b. The tree has a maximal depth of 31. The visualization was created by selecting a subset of depths, namely 0, 3, 9, and 12, and plotting all bounding volumes from nodes in the tree at this depth. Accordingly, figure 2.11a corresponds only to the bounding box of the brain compartment. We see that for higher depths, we get a tighter fit of the bounding volumes to the brain surface. Note furthermore that for a depth of 12 in figure 2.11d, it is quite noticeable that many triangles in fringe regions of the brain are no longer covered by a bounding volume, since the corresponding branches of the bounding volume hierarchy have already terminated in leaves of below 100 triangles at a lower depth.

This finishes our discussion on efficiently implementing the enforcement of a distance condition during the source space creation into DUNEuro. Before continuing, we want to note that we also implemented the enforcement of the Venant condition into DUNEuro. To this end, we extended the KDTree implementation in DUNEuro to support nearest neighbor queries, which, contrary to what is claimed in [186], was not implemented before.²⁷

As a final investigation in this section, we want to investigate the impact of enforcing a distance

²⁷In the previous implementation, it was only possible to, given a set of positions, compute a point from this set that is probably relatively close to the query point, which in combination with edgehopping was sufficient for a quick localization of the element containing the reference position. For details, please look at the Git history under https://gitlab.dune-project.org/duneuro/duneuro/-/blob/master/duneuro/common/kdtree.hh?ref_type=heads

condition on the generated source spaces. To this end, we constructed three different source spaces, say \mathcal{S}_1 , \mathcal{S}_2 , and \mathcal{S}_3 . Each of these source spaces is based on the same regular 3d grid with a resolution of 1.5 mm placed upon the realistic head model from figure 2.3b. For \mathcal{S}_1 , we only rejected candidate positions if they were not contained in the gray matter. For \mathcal{S}_2 , we rejected positions if they were either not contained in the gray matter or did not fulfill the Venant condition, and for \mathcal{S}_3 we rejected positions if they were either not contained in the gray matter or did not have a distance of at least 0.35 mm from the gray matter boundary. Then, we computed for each position in these source spaces the distance from the gray matter boundary, and plotted the resulting distribution in a histogram. The results can be seen in figures 2.12, 2.13, and 2.14.

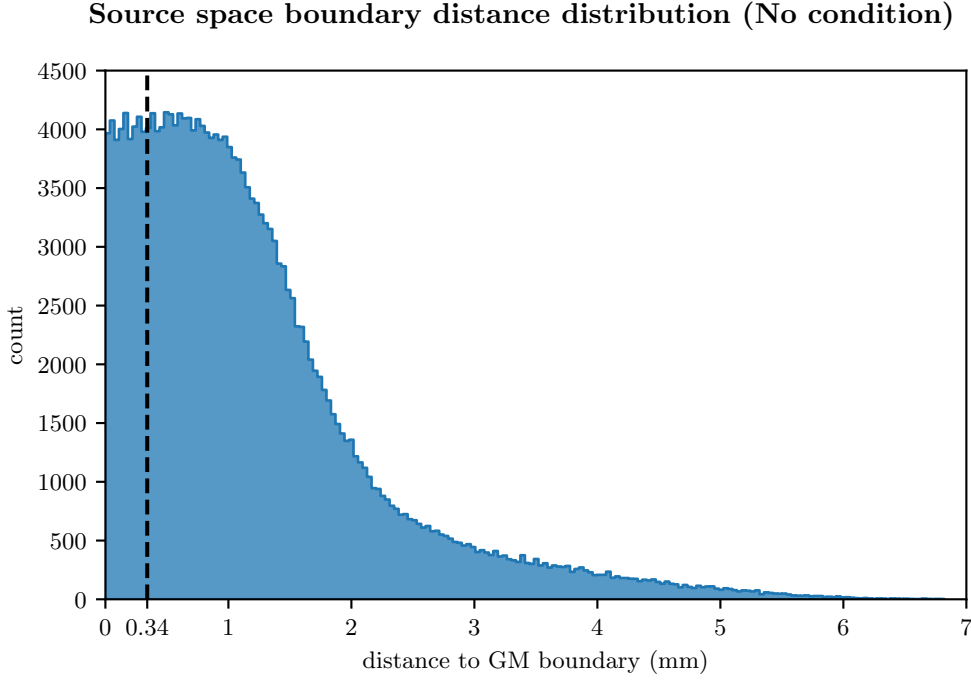


Figure 2.12: Distance distribution of a source space where it was only enforced that sources lie inside the gray matter compartment. Using a resolution of 1.5 mm, this yields a total of 214440 source positions. The distance of 0.34 mm, below which the numerical integration of the MEG local subtraction integrals becomes critical for elements with an edge size of 2 mm, is highlighted.

First of all, we see in figure 2.12 that only checking for containment in the gray matter produces a large number of sources that are very close to the compartment boundary, and we see that some form of additional rejection criterion is necessary. When looking at figure 2.13, we see that the Venant condition indeed massively cuts down on the number of sources close to a conductivity jump. But we also see that many sources at a reasonable distance, such as between 1 mm and 2 mm, are rejected, and there remains a relatively small subset of sources that are relatively close to a conductivity jump, for the same reason as shown in figure 2.10. In fact, in this concrete example, the Venant condition allowed for a source that has a distance of only 0.00091 mm to a conductivity jump.

With this in mind, we argue that strictly enforcing a distance condition, such as shown in

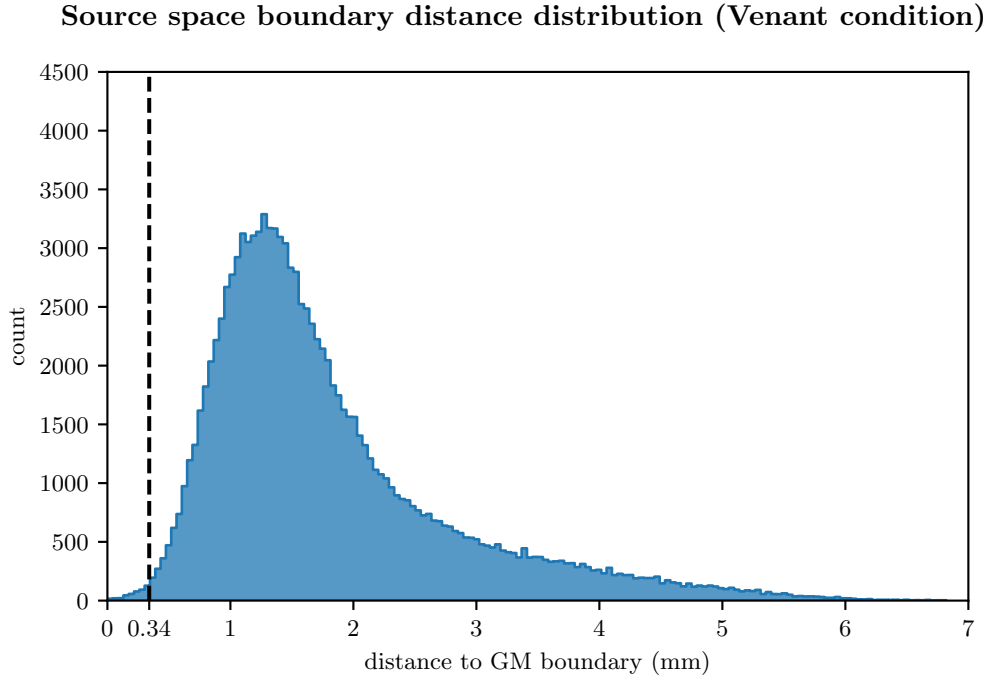


Figure 2.13: Distance distribution of a source space where it was enforced that sources lie inside the gray matter compartment and fulfill the Venant condition. Using a resolution of 1.5 mm, this yields a total of 113990 source positions. The distance of 0.34 mm, below which the numerical integration of the MEG local subtraction integrals becomes critical for elements with an edge size of 2 mm, is highlighted.

figure 2.14, is the superior way of constructing source spaces for the local subtraction approach. Furthermore, this investigation can also be interpreted as showing an advantage of the local subtraction approach over the different forms of the Venant approach. When constructing a source space for the Venant approach, one cannot really drop the Venant condition, as it is needed to guarantee that enough monopole positions are available to properly model the targeted dipolar behaviour, see section 2.5.1. But there might be positions in the mesh, e.g., in regions with a thin cortex, where it is not possible to place source positions fulfilling the Venant condition. If we, however, want to reconstruct neural activity from such a brain region, it has to be represented in the source space. If one wants to employ a Venant-type approach but does not want to risk a deterioration of the numerical accuracy, one essentially has to locally refine the FEM mesh in the relevant regions. The local subtraction approach, in contrast, is less restrictive with regard to the local mesh structure, and, even in thin cortices, we can expect an accurate forward simulation, given that we assemble the MEG integrals in a sufficiently accurate way.

2.7.4 Setting up a continuous integration pipeline

At the IBB, we develop a C++ toolbox for numerically simulating the electromagnetic fields generated by either neural activity or external stimulating devices. This toolbox is named DUNEuro [186], and builds on top of the DUNE framework [14]. While this toolbox is both highly efficient and supports a wide range of features, it also comes with a certain level of technical difficulty,

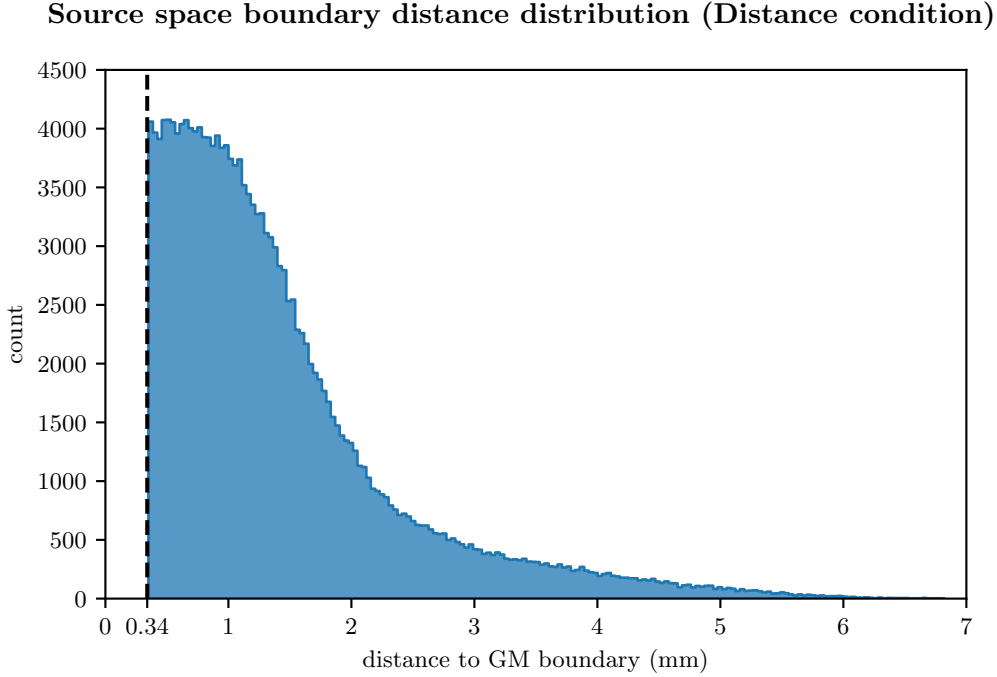


Figure 2.14: Distance distribution of a source space where it was enforced that sources lie inside the gray matter compartment and have a distance of at least 0.35 mm from a conductivity jump. Using a resolution of 1.5 mm, this yields a total of 176154 source positions. The distance of 0.34 mm, below which the numerical integration of the MEG local subtraction integrals becomes critical for elements with an edge size of 2 mm, is highlighted.

both with respect to development as well as with respect to usage. From the development side, DUNEuro makes heavy use of “advanced” programming features, such as templates, virtual functions, or SFINAE. And while these concepts are certainly not impossible to learn, they form a barrier that needs to be overcome before someone can contribute to the development of DUNEuro. Given the constraints of academic life, it is oftentimes not possible to build up extensive experience in the relevant coding practices, and one has to make do with a limited amount of understanding. Furthermore, even if someone is already familiar with C++ programming, getting familiar with a large toolbox and understanding its interdependencies remains challenging. From this perspective, it seems imperative that one should set up a system that automatically checks and validates proposed code changes to make sure that no existing functionality is broken, and give the implementer early feedback if they fail to meet certain expectations on the way the code is written. If one approaches DUNEuro from a user perspective instead, one also encounters a number of barriers. First of all, if one does not use an existing integration, such as the one into Brainstorm [129], the users currently have to compile the DUNEuro library by themselves. Anyone who has ever tried compiling DUNEuro knows that this can, and does, fail quite easily if one goes outside of one’s familiar compilation environment. And while one can usually apply some easy fixes to get the compilation working again, in my experience, users, especially those who have no prior experience in reading C++ compiler error messages, are, understandably, not willing to spend the time to figure out how to get their compilations to work. Furthermore, even if the compilation does work, a potential user has to figure out how to use the toolbox. This can

e.g. be achieved by providing example scripts, as e.g. done in [186]. Note, however, that while [186] is an excellent starting point for potential users, it needs to be complemented by additional information. First, the DUNEuro toolbox is continuously developed, and thus, over time, additional features will be added, which are then not covered by the example scripts supplied in the paper. Secondly, it might at some point be necessary to make changes to either the API or the installation process of the toolbox. This can lead to the code base changing in such a way that backwards compatibility is broken, which might render the published scripts from [186] deprecated.²⁸ Thus, we argue that it is essential to provide example scripts that demonstrate how the core set of functionality can be achieved in the current release of the toolbox.

We want to demonstrate that the aforementioned problems can be addressed, at least to some degree, by the introduction and maintenance of a *continuous integration (CI) pipeline*. The general principle goes as follows. If one collaboratively works on some software, one typically uses a version control software, such as Git. Then, there is typically some “central” version of the repository, from where contributors push and pull. If a contributor now pushes a change to this repository, one can automatically test if the pushed code still compiles and runs as expected before allowing the changes to be merged into the main branch.

There are different realizations of this concept. As the “central” DUNEuro repository is hosted as a GitLab instance²⁹, we will focus on GitLab in particular. Here, one can register a so-called *runner* to a repository. This runner can be thought of as a computer that monitors the repository and can react to events, such as a commit from a contributor. The runner can then fetch the committed changes and perform a set of specified operations. In GitLab, this set of operations is specified in a dedicated configuration file,³⁰. Concretely, this file specifies a collection of so-called *jobs*. Generally, if each job that is specified to run in the configuration file successfully completes, the pipeline is said to have *passed*, and if at least one job fails the pipeline is said to have *failed*.³¹

A job is, in essence, given by

- An initial container image in which commands are executed.
- A set of environment variables that are available inside the container.
- A list of shell instructions to execute inside the container.

The *container image*, such as, e.g., a Docker container image, allows the runner to simulate the running of an operating system that is different from the host system. While the runner might e.g. run on Ubuntu 22.04 as an operating system, containers enable the runner to e.g. simulate a “fresh” Ubuntu 24.04 operating system. One great benefit of containers is that they are in a known and reproducible state.

Imagine that a contributor has implemented a new method into the codebase that you want to test. But when trying to set up the experiment, you notice that the code fails to compile. In personal communication, it then turns out that the contributor forgot to communicate that an additional dependency was introduced, and that you need to set a specific compilation flag. After following these hints, you actually manage to run the code, but get different results than what

²⁸<https://xkcd.com/1172/>

²⁹<https://gitlab.dune-project.org/duneuro/duneuro>

³⁰see <https://docs.gitlab.com/ci/yaml/> for the structure of this file, and https://gitlab.dune-project.org/duneuro/duneuro/-/blob/master/.gitlab-ci.yml?ref_type=heads for a concrete example of how this is implemented in DUNEuro

³¹Note that the CI syntax allows the user to especially allow certain jobs to fail, and a pipeline might hence pass even if not all jobs successfully complete.

was reported by the contributor. After some further communication, you are told that you, of course, need to set this parameter to that value. Complying with the new instructions, you are finally able to reproduce the results.

If one enforces an adherence to a CI pipeline, situations such as the one sketched above would be avoided, since it forces the contributor to disclose, at the time of committing,

1. the operating system and its initial state that was used.
2. the additional steps needed to prepare the compilation.
3. the concrete flags and commands used for the compilation.
4. (ideally) a test of their newly implemented code that validates the output computed inside the container against a set of precomputed reference values.

The CI pipeline we created for DUNEuro consists, at the time of writing, of a total of 28 jobs. These jobs are organized into three stages. In a first stage, we check for spelling mistakes and licensing information. In a second stage, we test, for three different operating systems (Ubuntu 22.04, Ubuntu 24.04, and Debian 12), if the DUNEuro library still compiles and all implemented tests produce the expected results. For each of these operating systems, we perform this check once without the TBB library and once with the TBB library.³² Furthermore, we always check the compilation with two different toolchains, one based on the gcc compiler and one based on the clang compiler. If all of these jobs succeeded, a third stage is entered. In this stage, the bindings of DUNEuro are tested. DUNEuro offers interfaces to both Python and Matlab. For the operating systems and toolchains mentioned above, the third stage then tests if these bindings are still able to compile, and if calling the core functionalities through these interfaces still produces the expected results. Figure 2.15 illustrates what it looks like if this pipeline is passed.

For each job in the “build_and_test” stage, it is first validated that DUNEuro can be compiled from one of the default DUNE Docker images (see <https://gitlab.dune-project.org/docker/ci>) by the recipe specified in the configuration file. Then, a number of tests are run. Before we started work on the CI pipeline, three tests were implemented, of which two did not fail. We fixed the failing test and implemented 17 further tests, which we argue cover the essential functionality of DUNEuro. A list of these tests from one of the jobs in figure 2.15 is shown in figure 2.16.

For the jobs in the “bindings” stage, we validate the compilation with the given recipe, but, at the point of writing, only run 4 tests covering the basic calls to solve the forward problem for EEG and MEG with a direct and a transfer matrix approach.

This covers the current state of the DUNEuro CI pipeline. For future development, there are numerous possible ideas. First of all, the number of tests and the amount of code covered by them, especially for the Python and Matlab bindings, should be increased. One aspect that we, in my opinion, should place a high emphasis on is the creation and maintenance of a set of Docker images that contain a compiled and properly set up version of DUNEuro. I have repeatedly experienced that users who were initially interested in using DUNEuro abandoned this pursuit due to difficulties with the compilation process. If we were to offer the possibility to just pull a Docker image and, e.g., access DUNEuro through a Jupyter notebook in the browser, this might eliminate a major source of frustration among potential users and lead to a more widespread adoption of FEM methodology in the MEEG community. In the CI pipeline, we already build

³²TBB [221] is a library for parallel programming in C++, and is currently an optional dependency of DUNEuro.

implement local operator for analytical formulas for anisotropic sigma infinity

✓ Passed Malte Höltershinken created pipeline for commit [a6362834](#) 2 months ago, finished 2 months ago

1 related merge request: [!88 Draft: New release ... collection of new feature \(localized assembler, tdcs, source space creation\)](#)

branch 28 jobs 146 minutes 34 seconds, queued for 2 seconds

Pipeline Jobs 28 Tests 0

Group jobs by Stage Job dependencies

.pre	build_and_test	bindings
✓ code-spelling-check	✓ debian:12_clang-14-20	✓ debian:12_clang-14-20_matlab_bindings
✓ reuse	✓ debian:12_clang-14-20_noTBB	✓ debian:12_clang-14-20_python_bindings
	✓ debian:12_gcc-12-20	✓ debian:12_gcc-12-20_matlab_bindings
	✓ debian:12_gcc-12-20_noTBB	✓ debian:12_gcc-12-20_python_bindings
	✓ ubuntu:22.04_clang-14-20	✓ ubuntu:22.04_clang-14-20_matlab_bindings
	✓ ubuntu:22.04_clang-14-20_noTBB	✓ ubuntu:22.04_clang-14-20_python_bindings
	✓ ubuntu:22.04_gcc-11-20	✓ ubuntu:22.04_gcc-11-20_matlab_bindings
	✓ ubuntu:22.04_gcc-11-20_noTBB	✓ ubuntu:22.04_gcc-11-20_python_bindings
	✓ ubuntu:24.04_clang-18-23	✓ ubuntu:24.04_clang-18-23_matlab_bindings
	✓ ubuntu:24.04_clang-18-23_noTBB	✓ ubuntu:24.04_clang-18-23_python_bindings
	✓ ubuntu:24.04_gcc-14-23	✓ ubuntu:24.04_gcc-14-23_matlab_bindings
	✓ ubuntu:24.04_gcc-14-23_noTBB	✓ ubuntu:24.04_gcc-14-23_python_bindings

Figure 2.15: Screenshot of a passed CI pipeline on the DUNEuro GitLab server. At the top of the screenshot, you can see the message of the commit that triggered the CI pipeline. Each green check mark corresponds to a job that was successfully completed. If you have sufficient rights on the server, you can directly access the pipeline under <https://gitlab.dune-project.org/duneuro/duneuro/-/pipelines/77386>.

DUNEuro and its bindings inside a Docker container. Furthermore, we can make containers available through the container registry integrated into the DUNEuro GitLab server.³³ Given this, I think that offering DUNEuro Docker images is within the reach of our capabilities.

As a closing remark of this section, I want to stress that I think that maintaining the CI pipeline is of major importance. As outlined above, it is a massive aid in ensuring reproducibility of results, onboarding new contributors, and sharing our technically intricate mathematical developments with the wider bioelectromagnetism community. As such, I believe that investing resources into developing and maintaining the CI pipeline is a valuable investment in the future.

³³https://gitlab.dune-project.org/duneuro/duneuro/container_registry

```

2544 1/20 Test #1: test_electrode_projection ..... Passed 0.59 sec
2545 2/20 Test #2: test_numerical_flux ..... Passed 0.59 sec
2546 3/20 Test #3: test_physical_flux ..... Passed 0.59 sec
2547 4/20 Test #5: test_analytical_expressions_local_subtraction ..... Passed 0.58 sec
2548 5/20 Test #6: test_anisotropic_analytical_expressions_local_subtraction ... Passed 0.57 sec
2549 6/20 Test #4: test_bounding_volume_hierarchy ..... Passed 13.20 sec
2550 7/20 Test #7: test_eeg_cg_local_subtraction ..... Passed 25.07 sec
2551 8/20 Test #8: test_eeg_cg_subtraction ..... Passed 30.18 sec
2552 9/20 Test #10: test_eeg_cg_partial_integration ..... Passed 31.10 sec
2553 10/20 Test #9: test_eeg_cg_multipolar_venant ..... Passed 38.16 sec
2554 11/20 Test #17: test_eeg_dg_partial_integration ..... Passed 96.27 sec
2555 12/20 Test #16: test_eeg_dg_subtraction ..... Passed 98.21 sec
2556 13/20 Test #15: test_eeg_dg_local_subtraction ..... Passed 107.88 sec
2557 14/20 Test #19: test_eeg_transfer ..... Passed 108.23 sec
2558 15/20 Test #11: test_meg_cg_local_subtraction ..... Passed 160.17 sec
2559 16/20 Test #14: test_meg_cg_partial_integration ..... Passed 160.42 sec
2560 17/20 Test #13: test_meg_cg_multipolar_venant ..... Passed 163.98 sec
2561 18/20 Test #12: test_meg_cg_subtraction ..... Passed 175.55 sec
2562 19/20 Test #18: test_meg_dg_partial_integration_physical_flux ..... Passed 206.33 sec
2563 20/20 Test #20: test_meg_transfer ..... Passed 330.86 sec
2564 100% tests passed, 0 tests failed out of 20
2565 Total Test time (real) = 331.44 sec
2566 JUnit report for CTest results written to /builds/duneuro/duneuro/junit/duneuro-cmake.xml
2567 --- duneuro done ---
2568 --- done ---
2569 + [[ -v install_result ]]
2570 + [[ -n '' ]]
2571 + exit 0
✓ 2572 Cleaning up project directory and file based variables
2573 Job succeeded

```

Figure 2.16: List of the tests that were run inside the job “debian:12_clang-14-20” of figure 2.15.

2.8 Numerical investigations

After having derived our new local subtraction approach and discussed various mathematical aspects of its implementation, we now want to investigate its performance numerically. In section 2.8.2, we will first fix the parameters of the patch construction. In sections 2.8.3 and 2.8.4, we will then compare the local subtraction approach, the classical subtraction approach, and the state-of-the-art direct approaches in spherical head models, where analytical solutions are available. In section 2.8.5, we will finally investigate the impact of the local subtraction approach in realistic head models.

The following investigations largely follow our publication [91]. To aid the reproducibility of our results, we have made all the code and data that are needed to run the following tests openly available on Zenodo [96, 90, 97], and were awarded the “SIAM Reproducibility Badge”. Furthermore, we have prepared a Docker container based on Ubuntu 24.04 that includes both this data as well as a compiled version of the DUNEuro library and the scripts needed to reproduce the published results. In fact, if you have the Docker command line interface set up on your machine, you can just copy and paste the command

```
docker run -it --rm registry.dune-project.org/duneuro/duneuro/local_subtraction_sisc_code_ubuntu:24.04
```

into your terminal and reproduce the results *right now*.

2.8.1 Volume conductor models used for validation

We want to compare the different forward approaches in spherical and realistic head models. It is, however, difficult to ensure that such a comparison is “fair”. As different approaches are based on different numerical constructions, they also suffer and benefit from different properties of the underlying models. This enables investigators to skew the results by simply building the models in such a way that their method of choice is favoured.

Spherical volume conductor models

We want to try to be fair to the individual approaches. The “appropriate” local mesh resolution depends on the function we want to approximate. In the discussion following definition 2.5.10 of the EEG subtraction approach, we saw that the correction potential of the classical subtraction approach can be interpreted as arising due to charge accumulation at conductivity jumps. From a heuristic perspective, we would also expect that, especially at tissue interfaces, the true head potential u might strongly deviate from the homogeneous potential u^∞ , as, e.g., a large conductivity difference will strongly influence the induced currents. Given this, one expects that subtraction approaches need a sufficient resolution at tissue interfaces to achieve an accurate simulation. As both the Venant and Whitney approaches rely on placing monopolar loads on mesh vertices,³⁴ they require that sufficiently many mesh vertices are present in the proximity of the target dipole position to faithfully recreate the dipolar behaviour. Thus, these direct approaches are expected to perform well if there is a sufficient resolution of the brain compartment. These considerations are in line with the results from [116], where it is demonstrated that meshes that ensure a high resolution even in outer head regions, that are comparatively far away from the dipole position, are beneficial for the classical subtraction approach, while meshes that have a higher resolution in the brain compartment benefit the Venant and the partial integration approach.

The local subtraction approach, at least if one chooses a small patch and transition region, leads to a sparse right-hand side. Thus, similar to the direct approaches and in contrast to the classical subtraction approach, we expect the local subtraction approach to be less dependent on the resolution of the mesh in regions that are at a large distance from the source positions, such as the skin compartment. The local dipolar behaviour, on the other hand, is already baked into the homogeneous potential u^∞ , and thus does not need to be recreated from the local mesh structure. Thus, similar to the classical subtraction approach and in contrast to the direct approaches, we also expect the local subtraction approach to be less dependent on the local brain resolution. We thus expect that the local subtraction approach is accurate as long as the conductivity jump interfaces contained in the patch and transition regions have a sufficiently high resolution.

We want to investigate scenarios that are beneficial to each of the different approaches. Our underlying geometry is that of a multilayer sphere consisting of 4 layers. The parameters of this sphere are listed in table 2.2.

Using the Gmsh software [72], we then created an initial mesh, called *mesh_init*, with a high resolution in the CSF and skull compartments. We then used the UG software [13] to create two further meshes from this initial mesh. First, we refined the brain compartment of *mesh_init*, which renders the mesh more favourable to the direct approaches, yielding a mesh we called

³⁴Remember that, even though the Whitney approach is framed in an $H(\text{div}; \Omega)$ context, the actual functions used in the FEM forward simulation are equivalent to placing monopolar loads on two mesh vertices lying opposite from one another. See section 2.5.1 for details.

Table 2.2: Multi-layer sphere parameters. The table is taken from [91].

Compartment	Conductivity	Radius
Brain	0.33 S m^{-1}	0 - 78 mm
CSF	1.79 S m^{-1}	78 - 80 mm
Skull	0.01 S m^{-1}	80 - 86 mm
Skin	0.43 S m^{-1}	86 - 92 mm

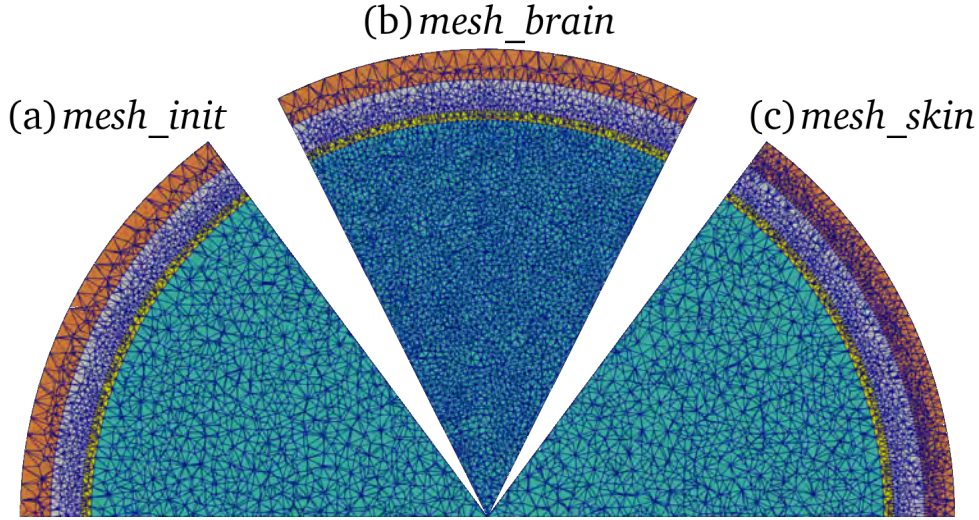


Figure 2.17: Circular sections of the spherical meshes used in the numerical tests. (a) shows the initial mesh with a high concentration of nodes in the CSF and skull compartment. (b) shows the mesh from (a) after refining the brain compartment. (c) shows the mesh from (a) after refining the skin compartment. Mesh (a) consists of ca. 800k nodes, mesh (b) consists of ca. 1.8 million nodes, and mesh (c) consists of ca. 1.3 million nodes. This figure, and its caption, is taken from [91].

mesh_brain. Then, we refined the skin compartment of *mesh_init*, which renders the mesh more favourable to the classical subtraction approach, yielding a mesh we called *mesh_skin*. The resulting meshes are visualized in figure 2.17.

To simulate EEG and MEG forward solutions, we also need to specify a set of measurement sensors. To this end, we first generated 200 approximately uniformly distributed positions on the outer surface of the multilayer sphere, which were then projected to the closest node position in *mesh_init*. These projected positions were then taken as electrode positions for the EEG forward problem. We then constructed a further 256 positions at a distance of 110 mm to the multilayer sphere center, again approximately uniformly distributed over the corresponding sphere surface. We then took each of these positions as a coil position for the MEG forward problem. Furthermore, we computed the complete secondary magnetic field vector for each of these coil positions, i.e., in this model, the output of the MEG forward problem for a single dipole is given by $3 \cdot 256 = 768$ measurement values.

Furthermore, we have already discussed in section 2.7.3 that the numerical simulation of EEG and

MEG measurements becomes more difficult the closer the source position is to a conductivity jump. Accordingly, we will do the same as essentially every other paper on EEG and MEG forward modeling and investigate how the errors evolve at different *eccentricities*. If the innermost shell of the multilayer sphere model has radius $r_0 > 0$, the sphere center is at $c \in \mathbb{R}^3$, and we have a dipole at position $x_0 \in B_{r_0}(c)$, we say that the source has an eccentricity of $\epsilon := \frac{\|x_0 - c\|_2}{r_0}$. Note that $0 \leq \epsilon < 1$. For the investigations in spherical models, we always decide on a selection of eccentricities we want to investigate. For each selected eccentricity ϵ , we then place 1000 sources approximately uniformly distributed on $\partial B_{\epsilon r_0}(c)$. For each of these sources, we then first numerically compute the EEG/MEG forward solution using the approach currently under investigation, and then compare the numerical results with the analytical solutions computed using either [39] in the EEG case or [178] in the MEG case. Furthermore, we want to note that in the EEG case we always use the *common average referenced* forward solution,³⁵ and, in the MEG case, we compare the numerical accuracies of the *secondary fields*, as the primary fields can be computed analytically, as can be seen in definition 2.1.3.

In the spherical model, we call a dipole $M \cdot \delta_{x_0}$ *radial* if M is a multiple of $x_0 - c$, where c denotes the sphere center, and we call the dipole *tangential* if M is orthogonal to $x_0 - c$. Since radial and tangential dipoles sometimes show different behaviour, we will investigate them separately. One concrete example is that, in spherical head models, the magnetic field generated by a radial dipole vanishes outside of the volume conductor [178]. In the MEG context, we will thus only investigate tangential sources in our numerical tests, while for the EEG case, we will investigate radial and tangential sources.

Finally, in [152] it is argued that, since the cortex has a thickness of 2 mm to 4 mm and equivalent current dipoles are located at the center of the cortical sheet, the sources at a distance of 1 mm to 2 mm to a conductivity jump are the ones that are most relevant for comparison with the realistic setting. When comparing different eccentricities, we will follow this consideration and highlight the eccentricities in this range.

Realistic volume conductor models

In realistic volume conductor models, we do not have access to analytical solutions, and thus have no “hard” criterion to compare one approach to the other. We can, however, give a qualitative view on how the approaches differ. We want to discuss the differences between the classical and the local subtraction approach by illustrating how the correction potentials they define differ in realistic head models, and in particular how they are influenced by conductivity jumps. The corresponding investigation utilizes the realistic volume conductor model from figure 2.1a. An additional visualization of this model can be seen in figure 2.18. This investigation largely follows the corresponding section from [91]. At the point of finishing up the revision of [91], our in-house pipeline was not able to handle anisotropic conductivities for tetrahedral meshes, and we thus used a default isotropic conductivity value for the white matter. We are aware that the presence of white matter anisotropy can have a significant effect on the generated electromagnetic fields [237], and this might be a topic for future investigations. For this investigation, however, we will restrict ourselves to the isotropic case, and we do not expect that our main takeaways about the difference between the local subtraction and the classical subtraction approaches will be significantly influenced by the presence of white matter anisotropy.

Concretely, we investigated the difference between subtraction approaches for a dipolar source

³⁵Remember that the EEG forward solution is only determined up to a constant, see e.g. corollary 2.4.6. If now $U_0 \in \mathbb{R}^L$ denotes an arbitrary forward solution vector, there exists a unique $c \in \mathbb{R}$ such that $U_0 + c \cdot 1$ is orthogonal to $1 \in \mathbb{R}^L$. $U_0 + c \cdot 1$ is then called the *common average referenced forward solution*.

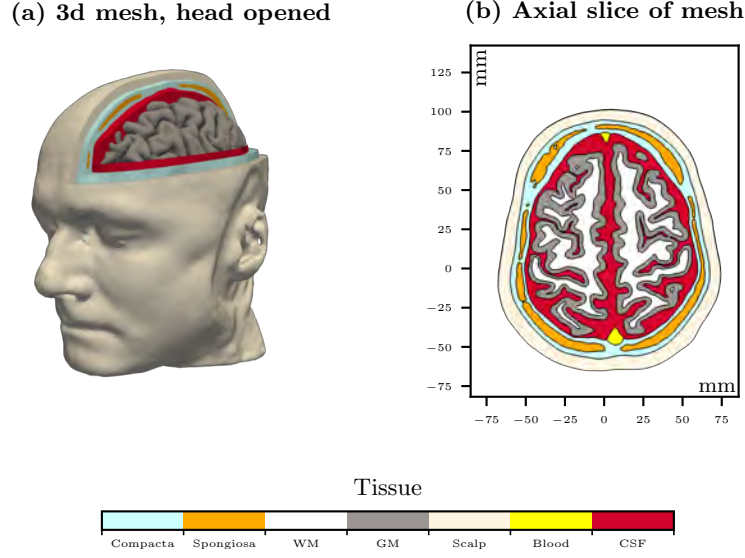


Figure 2.18: A further visualization of the realistic volume conductor model from figure 2.1a. (a) shows a 3d view of the volume conductor model, where we removed parts of the skin, skull, and CSF to visualize the embedding of the brain inside the head. Note that the opening of the head was done only for the visualization and not for the numerical simulation. (b) shows an axial slice through the model. On this slice, we sampled the conductivity with a resolution of 0.1 mm. From the resulting image, we first extracted boundaries and then dilated them to increase visibility. The corresponding image processing was performed in Python using the scikit-image library [208]. The colorbar at the bottom shows how colors are mapped to compartments and applies to (a) and (b). This figure, and its caption, are taken from [91].

modeling the response in the brain due to a somatosensory stimulus. This somatosensory brain response is of particular importance for our workflow at the IBB, as we typically use it to calibrate skull conductivity values [185, 6]. If one applies a somatosensory stimulation a few hundred times and then averages the resulting EEG and MEG measurements, one sees a peak at ca. 20 ms after the stimulus is applied, whose generating neural activity is typically attributed to the posterior bank of the central sulcus [105]. Furthermore, it is well known, and also obvious by looking at topographic plots of the measurements, that the current generating this so-called *P20 component* is flowing roughly tangential to the head surface. Based on this prior information, in combination with a dipole scan applied to the measurement data from [154], we estimated the location of the somatosensory response in the model shown in figure 2.18 and placed a dipolar source of strength 20 nAm at this position.³⁶

2.8.2 The local subtraction patch construction

In section 2.7.1, we proposed to construct the patch Ω^∞ and the transition region $\tilde{\Omega}$ by what we called *vertex extensions*, and then construct χ as a FEM trial function. At the end of that subsection, it was not yet clear how many patch extensions should be performed to achieve the

³⁶If the reader is interested in how precisely the measurement data from somatosensory stimuli is integrated into our workflow, we encourage them to look at section 4.2.1, where the head model creation for an epilepsy patient is described.

best outcome. This is what we want to investigate here.

To this end, we took the meshes from figure 2.17 and 1000 dipoles at an eccentricity of 0.99 and computed, for different numbers of vertex extensions, the errors against the analytical solution both for EEG as well as for MEG. As the classical subtraction approach is the local subtraction approach that arises when enough vertex extensions are performed to cover the whole head domain, it can be thought of as a “limit” of the local subtraction approach, and we hence also computed it for reference. In the following, we will only show the results for tangential dipoles, as radial dipoles show qualitatively the same results and only differ in the absolute values of the errors.

In the EEG case, figure 2.19 shows the result for `mesh_init`, figure 2.20 shows the result for `mesh_brain`, and figure 2.21 shows the result for `mesh_skin`. We see that for all three meshes, the error initially decreases up until 2 vertex extensions. Then, the errors do not change for a while, until they, maybe surprisingly, go up again. Furthermore, it is visible that as the number of vertex extensions increases, the error plots “converge” towards the classical subtraction errors, as one would expect.

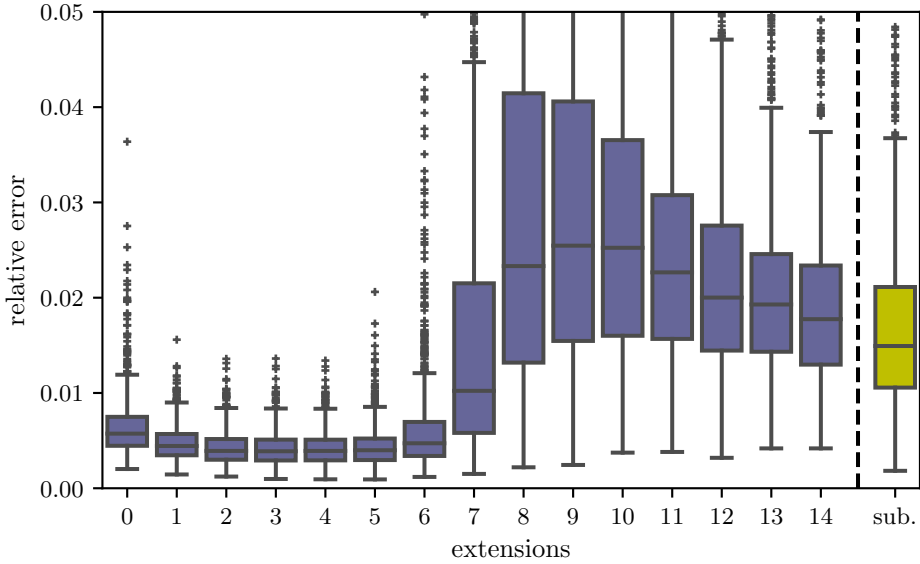


Figure 2.19: Relative error of the local subtraction approach in the EEG case for 1000 tangential dipoles at 0.99 eccentricity for different numbers of vertex extensions during patch construction, computed using `mesh_init` from figure 2.17. The rightmost yellow boxplot shows the errors when employing the classical subtraction approach using an analytical computation of the RHS, as described in [18]. This figure, and its caption, were taken from [91] and slightly modified.

The behaviour of these plots can be explained with the considerations from section 2.7.1. Initially, choosing a larger patch “subtracts” more of the singular dipole behaviour out of the problem, so that the correction potential becomes more regular and can thus be better approximated. But, as we expect a large deviation of u from u^∞ at tissue boundaries, solving for $u^c = u - \chi \cdot u^\infty$ at these regions might be more difficult than directly solving for u . We think that this is the reason for the increase in the errors at ca. 6 extensions, as there, the patch starts to grow into the skin compartment of the corresponding meshes. Looking at figure 2.17, one sees that the skin compartment has a relatively low resolution in both `mesh_init` and `mesh_brain`, which we think

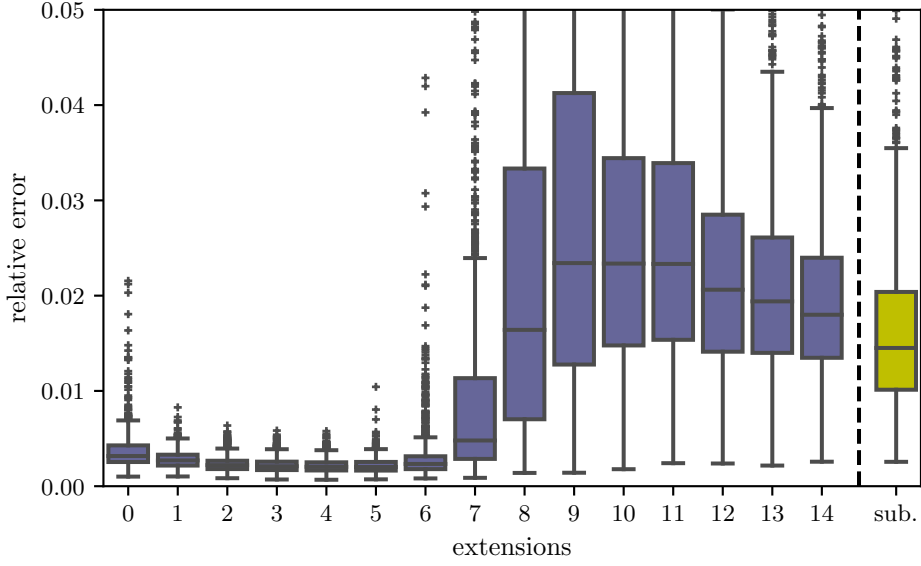


Figure 2.20: Relative error of the local subtraction approach in the EEG case for 1000 tangential dipoles at 0.99 eccentricity for different numbers of vertex extensions during patch construction, computed using `mesh_brain` from figure 2.17. The rightmost yellow boxplot shows the errors when employing the classical subtraction approach using an analytical computation of the RHS, as described in [18]. This figure, and its caption, were taken from [91] and slightly modified.

is insufficient to accurately resolve the difference $u - \chi \cdot u^\infty$, leading to the larger errors. Indeed, refining the skin compartment noticeably reduces the errors when performing a larger number of extensions, while barely changing the errors if one performs only a few extensions, as is seen in figure 2.21. Furthermore, if one plots the relative differences between the numerical solutions for small and large patches, one sees that the differences are largest in the elements adjacent to the skull-skin interface close to the source position. We think that this also explains the decrease that one e.g. sees in figure 2.19 when going from 9 to 14 extensions, as larger patches imply that we are approximating $u - u^\infty$, and not the more complicated $u - \chi \cdot u^\infty$, at the skull-skin interface close to the source.

To summarize, we see that larger patches generally lead to a larger dependence of the approach on the resolution of distant tissue interfaces. For the EEG problem, it thus seems that performing 2 vertex extensions strikes the optimal balance between building up a distance to the dipole singularity while simultaneously keeping the dependency on the global mesh structure minimal. As `mesh_brain` results from refining `mesh_init`, and in both figure 2.19 and figure 2.20 performing 2 vertex extensions seems to be optimal, we expect that this conclusion does not depend on the mesh resolution.³⁷

In the MEG case, the corresponding result for `mesh_init` is shown in figure 2.22. Similar to the EEG case, we also see an initial decrease for the first 2 vertex extensions. But in contrast to

³⁷We want to note that preliminary results from the master's thesis [79] suggest that, in certain meshes with a high resolution, it might be beneficial for the accuracy of the local subtraction approach to perform more than two vertex extensions. However, in these meshes, two vertex extensions also yield highly accurate simulation results, and, as a compromise between computational efficiency and accuracy, we argue that the recommendation of performing two vertex extensions is still sensible. Note that this assessment is based on a preliminary version of the master's thesis.

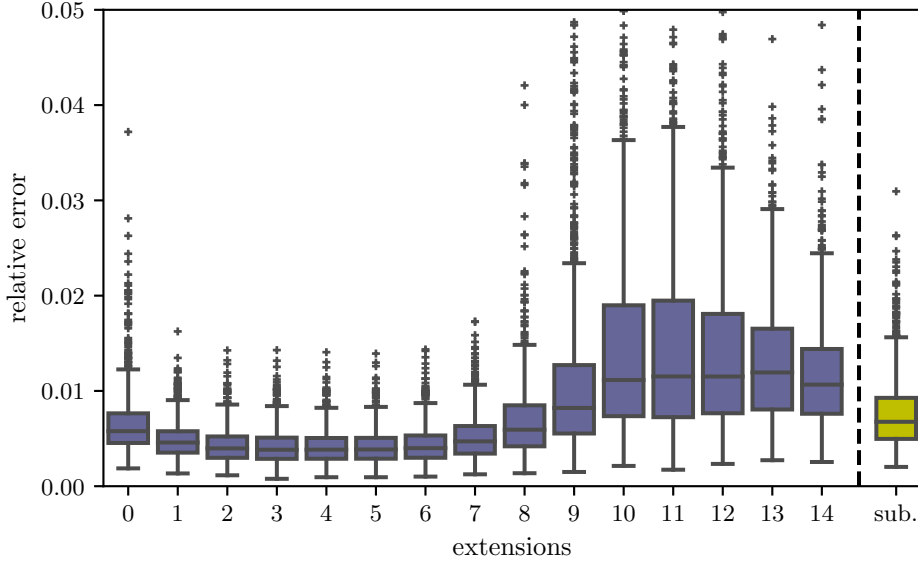


Figure 2.21: Relative error of the local subtraction approach in the EEG case for 1000 tangential dipoles at 0.99 eccentricity for different numbers of vertex extensions during patch construction, computed using `mesh_skin` from figure 2.17. The rightmost yellow boxplot shows the errors when employing the classical subtraction approach using an analytical computation of the RHS, as described in [18]. This figure, and its caption, were taken from [91] and slightly modified.

the EEG case, the errors do not increase when performing a large number of vertex extensions. The reason for this is that the errors that are visible in figure 2.19 for a large number of patch extensions are mainly located in the skin compartment. But the currents in the skin compartment only have a negligible contribution to the Biot-Savart integral, as, due to the insulation of the skull, the majority of currents are concentrated inside the brain and CSF compartments. We also performed the same experiment for `mesh_brain` and `mesh_skin`, and got the same qualitative result.

We thus arrive at the same conclusion as in the EEG case, i.e., that performing 2 vertex extensions for the patch construction seems optimal with regard to accuracy. Furthermore, from a computational perspective, it is also desirable to perform only two vertex extensions, as building the patch with a larger number of vertex extensions would also lead to a larger computational demand of the approach.

2.8.3 Comparing direct approaches for the forward problems

In section 2.5.1, we introduced the partial integration, the Whitney, and the Venant classes of approaches for the EEG and MEG forward problems. While it has been demonstrated that the Whitney approaches, the multipolar Venant approach, and the variant of the Venant approach from Andreas Nüking’s PhD thesis all perform better than both the partial integration approach as well as the original Venant approach due to Buchner et. al. [162, 218, 143], these approaches have, to the best of our knowledge, not been compared against one another. This comparison is the topic of this section.

We have already noted before that for the direct approaches, the main limiting factor for their

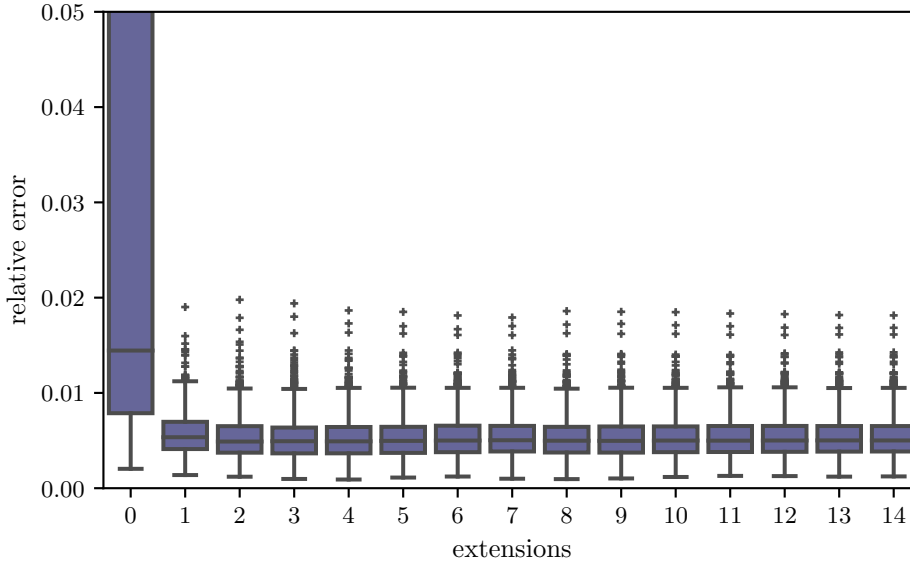


Figure 2.22: Relative error of the local subtraction approach in the MEG case for 1000 tangential dipoles at 0.99 eccentricity for different numbers of vertex extensions during patch construction, computed using `mesh_init` from figure 2.17. In contrast to, e.g., figure 2.19, we did not include the classical subtraction approach in the plot, as it would look exactly the same as the plot for 14 extensions.

accuracy is the number of mesh nodes in the proximity of the dipole position. Accordingly, all results we generated for `mesh_skin` were almost identical to the corresponding results for `mesh_init` from figure 2.17. We will, hence, in the following only discuss `mesh_init` and `mesh_brain`.

To not clutter the thesis with too many pictures, we will give a sample of the results that we argue is representative. In figure 2.23 the results for the EEG forward problem and radial dipoles in `mesh_init` are shown, in figure 2.24 the results for the EEG forward problem and radial dipoles in `mesh_brain` are shown, and in figure 2.25 the results for the MEG forward problem and tangential dipoles in `mesh_init` are shown.

First of all, we want to note that over all investigated combinations of mesh, dipole orientation, and modality, it was always the case that the median of the multipolar Venant approach was the lowest among the four competing direct approaches. Furthermore, looking at figure 2.23, we see a strong mesh dependency of the Whitney approaches. In `mesh_init`, sources with an eccentricity of 0.9639 tend to land on the inner boundary of the outer element row of the brain compartment, while sources with an eccentricity of 0.9778 tend to land more in the middle of the outer brain element row. This mesh dependency is not visible for the Venant approaches. This agrees with the findings in [218], where it is shown that the accuracy of the multipolar Venant approach is much less dependent on the location of the dipole position within an element, when compared to the original Buchner Venant. The only thing preventing us from giving a clear recommendation for the multipolar Venant approach over the other direct approaches is the large number of outliers it produces. Both for the EEG problem in figure 2.23 as well as for the MEG forward problem in figure 2.25, starting at an eccentricity of 0.9848, the multipolar Venant approach produces a large number of outliers. The reason for this is that at this eccentricity, and also at larger eccentricities, oftentimes the closest vertex to the dipole position happens to

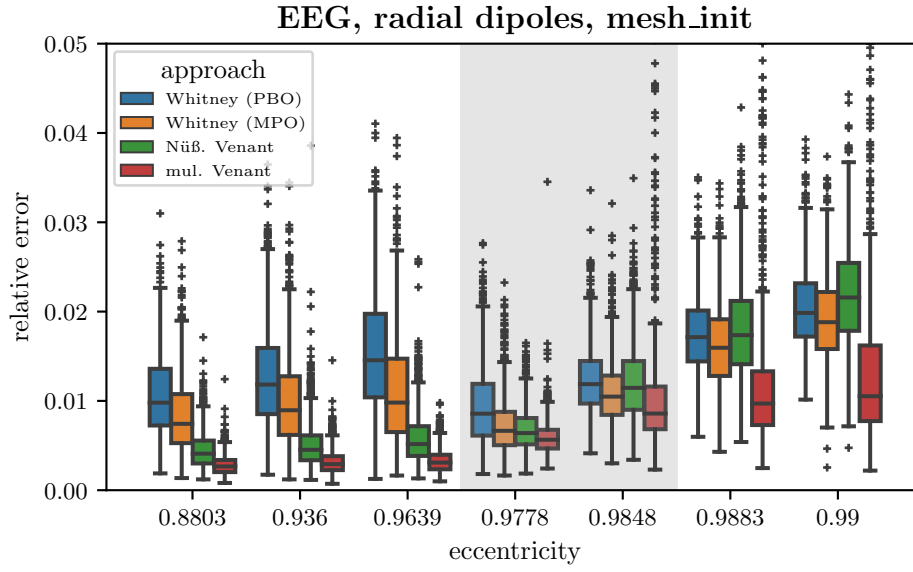


Figure 2.23: Accuracy comparison of different direct approaches for the EEG forward problem in mesh_init from figure 2.17. Eccentricities corresponding to a distance of 1 mm to 2 mm from the brain-CSF boundary are highlighted.

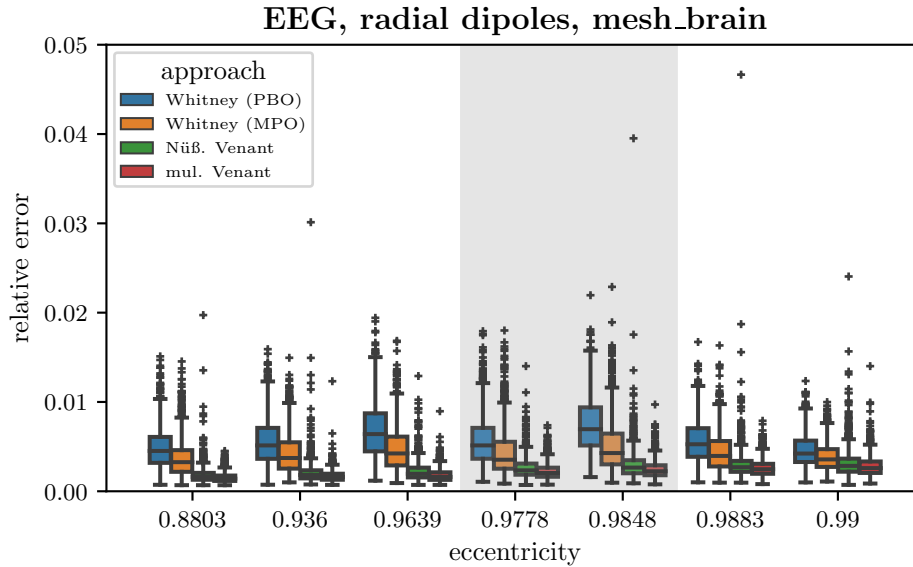


Figure 2.24: Accuracy comparison of different direct approaches for the EEG forward problem in mesh_brain from figure 2.17. Eccentricities corresponding to a distance of 1 mm to 2 mm from the brain-CSF boundary are highlighted.

lie on the brain-CSF boundary, i.e., these positions violate the Venant condition from section 2.7.3. What we see in these figures is thus the numerical instability arising from not adhering to the Venant condition. Furthermore, if we compare this with the results from figure 2.24, we see that the outliers almost entirely disappear, as refining the brain compartment introduces many

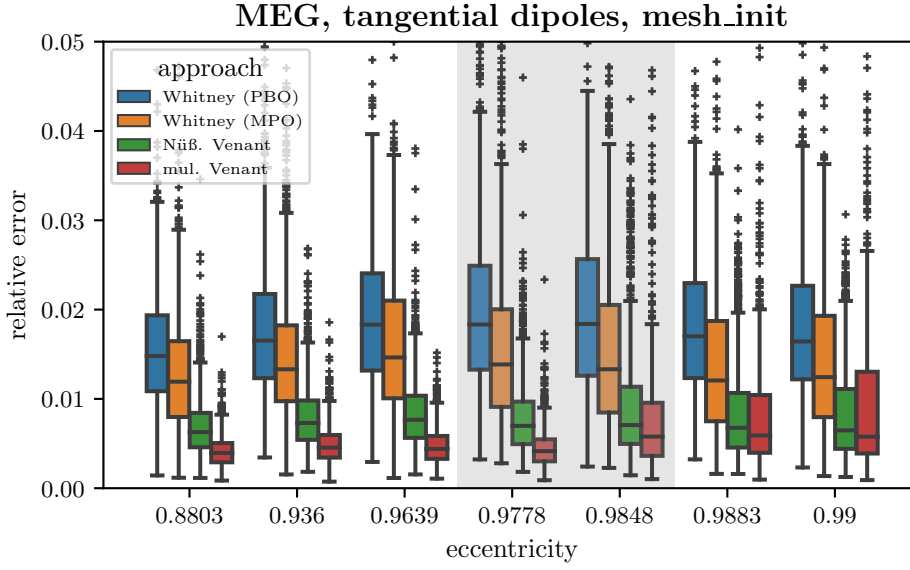


Figure 2.25: Accuracy comparison of different direct approaches for the MEG forward problem in `mesh_init` from figure 2.17. Eccentricities corresponding to a distance of 1 mm to 2 mm from the brain-CSF boundary are highlighted.

new brain vertices, and the Venant condition will hence be fulfilled much more often.

Our takeaway from this investigation is that the multipolar Venant approach is generally the most accurate direct forward approach in the published literature. If one wants to employ a direct approach, our recommendation for the practice would be to use the multipolar Venant approach in combination with a source space that fulfills the Venant condition.

2.8.4 Comparing direct and subtraction-type approaches

In this section, we want to compare the local subtraction approach with 2 vertex extensions (see section 2.8.2) against the classical subtraction approach and the multipolar Venant approach. We have shown in the preceding subsection that the multipolar Venant approach is generally the most accurate direct approach, and hence argue that a comparison against the remaining direct approaches is not really necessary.

In [91], we present all combinations of EEG/MEG, radial/tangential sources, and mesh selection from figure 2.17. Here, we will only present a subset of these that, in our opinion, suffice to illustrate the behaviour of the methods. If the reader is interested in the remaining comparisons, we thus refer to [91]. Figure 2.26 shows the results for the EEG forward problem, radial dipoles, and `mesh_init`, figure 2.27 shows the results for the MEG forward problem, tangential dipoles, and `mesh_init`, figure 2.28 shows the results for the EEG forward problem, radial dipoles, and `mesh_brain`, and figure 2.29 shows the results for the EEG forward problem, radial dipoles, and `mesh_skin`.

In figure 2.26, we see that, in `mesh_init` for the EEG forward problem, the local subtraction approach produces the best results out of the three investigated approaches. For the local subtraction approach, we see that the “box” of the boxplot stays below 1% over all investigated eccentricities, and, in contrast to the multipolar Venant approach, it produces no large outliers

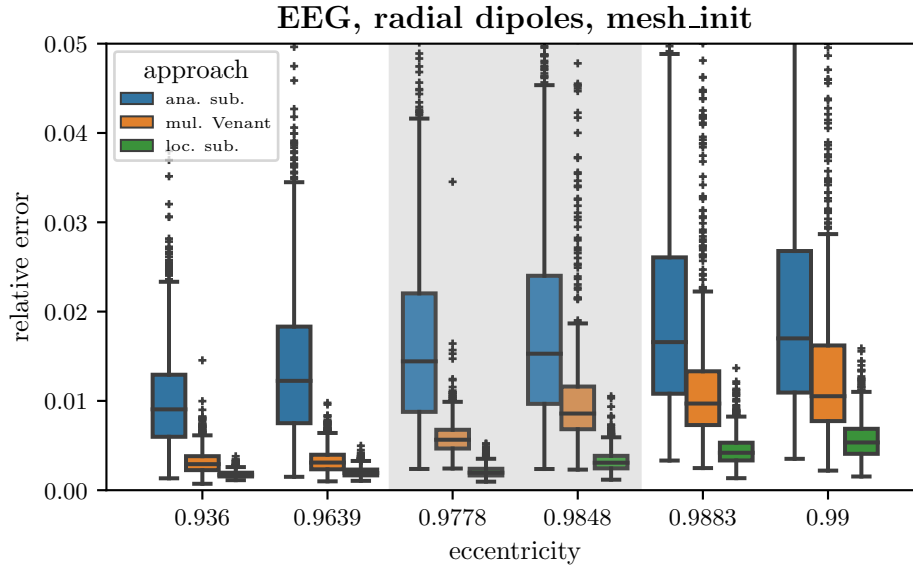


Figure 2.26: Accuracy comparison of the classical subtraction approach (abbreviated as ana. sub., following the notation in [18]), the multipolar Venant approach, and the local subtraction approach for the EEG forward problem in `mesh_init` from figure 2.17. Eccentricities corresponding to a distance of 1 mm to 2 mm from the brain-CSF boundary are highlighted. This figure is taken from [91].

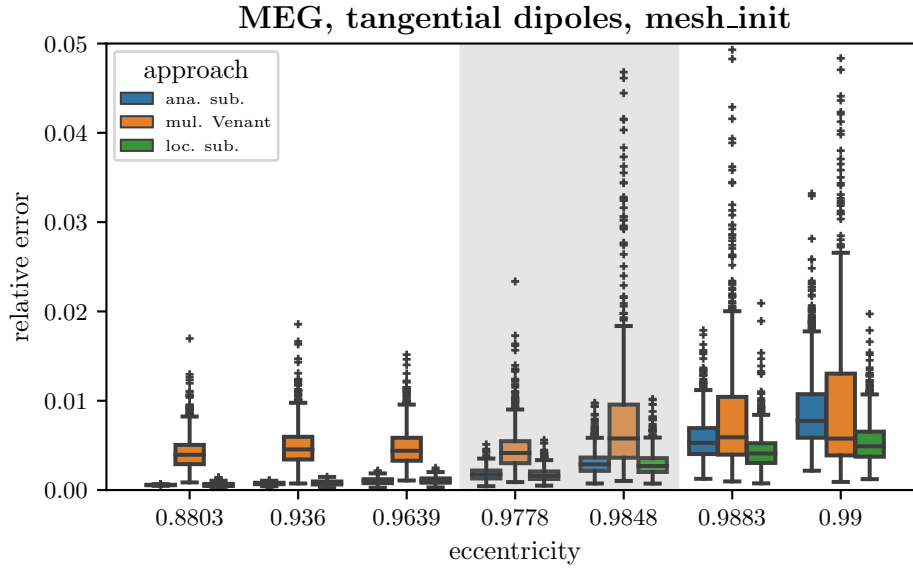


Figure 2.27: Accuracy comparison of the classical subtraction approach, the multipolar Venant approach, and the local subtraction approach for the MEG forward problem in `mesh_init` from figure 2.17. Eccentricities corresponding to a distance of 1 mm to 2 mm from the brain-CSF boundary are highlighted. This figure is taken from [91].

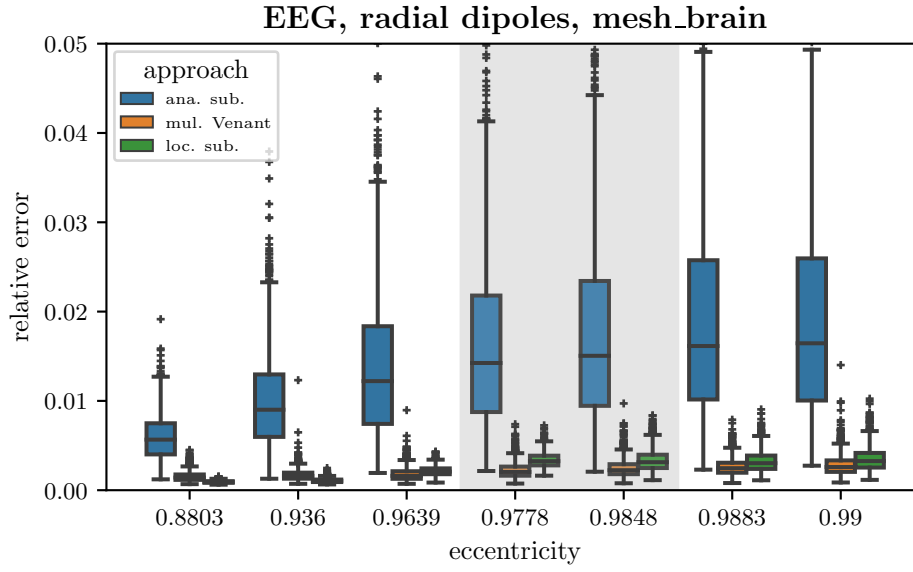


Figure 2.28: Accuracy comparison of the classical subtraction approach, the multipolar Venant approach, and the local subtraction approach for the EEG forward problem in mesh_brain from figure 2.17. Eccentricities corresponding to a distance of 1 mm to 2 mm from the brain-CSF boundary are highlighted. This figure is taken from [91].

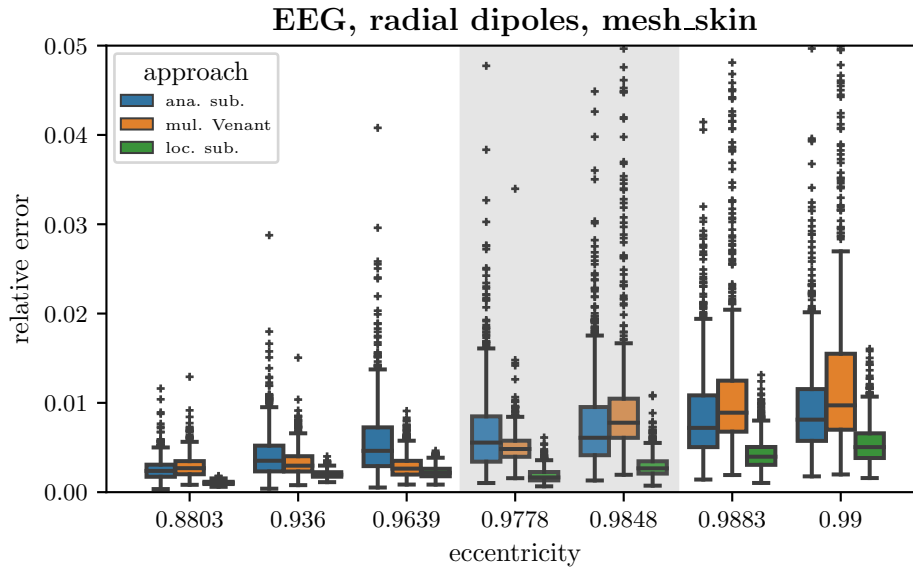


Figure 2.29: Accuracy comparison of the classical subtraction approach, the multipolar Venant approach, and the local subtraction approach for the EEG forward problem in mesh_skin from figure 2.17. Eccentricities corresponding to a distance of 1 mm to 2 mm from the brain-CSF boundary are highlighted. This figure is taken from [91].

even at high eccentricities. We also see that the classical subtraction approach produces relatively large errors over all investigated eccentricities.

These results are in line with our own results from the previous subsections. Due to the somewhat low brain resolution of `mesh_init`, it is difficult for the Venant approach to properly model the dipolar behaviour of the source. Furthermore, due to the low skin resolution, it becomes difficult for the classical subtraction approach to approximate $u^c = u - u^\infty$, which exhibits a non-trivial behaviour at tissue boundaries. Additionally, due to the insulating effect of the skull, we expect $\|\nabla u\| \ll \|\nabla u^\infty\|$ in the skin, and thus also $\nabla u^c \approx -\nabla u^\infty$, and computing $u = u^c + u^\infty$ can lead to cancellation effects reducing the accuracy of the approach.

If we, e.g., refine the brain compartment, we accordingly see in figure 2.28 that the errors for the multipolar Venant approach are strongly reduced, and it produces EEG results that are very comparable with the local subtraction approach, even slightly beating the local subtraction approach at some eccentricities. If we, however, perform the same experiment with tangential dipoles, we see that the multipolar Venant approach and the local subtraction approach are again very comparable, but now with a similarly slight edge for the local subtraction approach. This shows that, in meshes with a highly refined brain compartment, we can expect similar EEG accuracies for the local subtraction and the multipolar Venant approach. The errors for the classical subtraction approach, however, barely change, as refining the brain compartment does not change the resolution of the skin compartment, which, at least in this setting, is the limiting factor for the classical subtraction EEG accuracy.

If we instead refine the skin compartment, we see in figure 2.29 that, when compared to figure 2.26 for `mesh_init`, the errors for the multipolar Venant approach barely improve, as a higher resolution in the skin compartment do not enable a better modeling of the dipole in the brain compartment, which, in this setting, is the limiting factor for the multipolar Venant EEG accuracy. The errors of the classical subtraction approach, however, are greatly reduced, as a better resolution of the skin compartment enables a better approximation of $u - u^\infty$ there. But even in this scenario, the local subtraction still performs better than the classical subtraction over all investigated eccentricities.

The errors of the local subtraction approach, in comparison, barely change at all when refining either the brain or the skin compartment, except that the errors for the brain refined mesh at high eccentricities are slightly decreased. This shows that the local subtraction approach is much less dependent on the global mesh structure than the classical subtraction approach, but also less dependent on the local mesh structure than the multipolar Venant approach. Thus, while under optimal conditions the local subtraction approach and the multipolar Venant approach are both able to produce highly accurate results, the local subtraction approach is still able to produce accurate results even if the mesh locally has a non-optimal structure. Especially if we then think of a realistic head model, with its highly folded surfaces, we expect a greater numerical stability of the local subtraction approach, as there might not always be enough “good” mesh vertex positions around for the multipolar Venant approach to construct a faithful approximation of a dipole.

For the MEG forward problem, we see in figure 2.27 that in `mesh_init` both types of subtraction approaches outperform the multipolar Venant approach over all investigated eccentricities. We get the same qualitative result for `mesh_brain` and `mesh_skin`, with the errors being essentially identical in `mesh_skin` and a bit smaller in `mesh_brain`. When comparing the MEG results of figure 2.27 with the EEG results in figure 2.26, we get the, at first surprising, result that while the classical subtraction approach produces large errors in the EEG case, it produces highly accurate results in the MEG case.³⁸ This is due to the same reason as the behaviour seen in figure 2.22,

³⁸Note that for the EEG in figure 2.26 we have used *radial* sources, while for the MEG in figure 2.27 we have used *tangential* sources. But if we instead use tangential sources also for the EEG, the classical subtraction approach

i.e. that the currents close to the source have a comparatively high contribution to the Biot-Savart integral (resp. the Biot-Savart like terms we assemble, see definition 2.5.17.), while the currents in the skin only have a negligible contribution. As the errors of the classical subtraction approach are mainly located in the skin compartment, it can still produce an accurate MEG simulation, even though it leads to an inaccurate prediction of the electric potentials at the skin. We also want to emphasize that the rise in the MEG errors in figure 2.27 for the eccentricities 0.9883 and 0.99 is not an inherent property of the method, but merely a consequence of a “bad” implementation. For the (local) subtraction MEG problem, we need to numerically compute a set of integrals, see definition 2.5.17. These need to be evaluated, e.g., over the patch Ω^∞ . For the local subtraction approach, the computation was performed using the strategy derived in section 2.7.2, which, since the patch is small, is reasonably fast. For the classical subtraction approach, however, the patch is the whole head domain, i.e., $\Omega^\infty = \Omega$. In order to not wait too long, we decided to use an integration order of 5 for the computation of the MEG classical subtraction integrals.³⁹ If one were to use a higher integration order, the accuracy of the MEG classical subtraction approach in figure 2.27 would be as good as the accuracy of the MEG local subtraction approach. In total, we see a clear advantage of using subtraction-type approaches for the MEG forward problem. We expect that this is due to their ability to accurately model the local behaviour of the current density in the environment of the source position. For the direct approaches, this local behaviour, by construction, is somewhat different from that of a point dipole.

Finally, we also want to compare the computational demands of the different approaches. To this end, we measured the computation times of 1000 dipoles at an eccentricity of 0.99 on an AMD Ryzen Threadripper 3960X processor with 24 cores. The results can be seen in tables 2.3, 2.4, and 2.5. The time needed to set up the DUNEuro driver and compute the transfer matrices was not included in these time measurements.

Table 2.3: Forward simulation times for 1000 sources at an eccentricity of 0.99 in *mesh_init*. This table is taken from [91], except for the value for the MEG local subtraction approach, which was recomputed using the new formulation from definition 2.5.17.

Approach	loc. sub.	ana. sub.	mul. Venant
EEG time (s)	0.93	655	0.02
MEG time (s)	8.5	18925	0.04

Table 2.4: Forward simulation times for 1000 sources at an eccentricity of 0.99 in *mesh_brain*. This table is taken from [91], again except for the MEG local subtraction value that was recomputed.

Approach	loc. sub.	ana. sub.	mul. Venant
EEG time (s)	0.96	1396	0.02
MEG time (s)	4.0	43175	0.04

First of all, we see that the multipolar Venant approach is, by far, the fastest of the three approaches, enabling the EEG and MEG forward simulation of 1000 sources in below 100 ms. If

still produces much larger errors than both the multipolar Venant as well as the local subtraction approach, just as for the radial sources in figure 2.26. The corresponding figure can be found in [91].

³⁹Even with an integration order of 5, the computations underlying figure 2.27 took more than 2 full days on an AMD Ryzen Threadripper 3960X processor with 24 cores, essentially only due to the MEG classical subtraction approach.

Table 2.5: Forward simulation times for 1000 sources at an eccentricity of 0.99 in *mesh_skin*. This table is taken from [91], again except for the MEG local subtraction value that was recomputed.

Approach	loc. sub.	ana. sub.	mul. Venant
EEG time (s)	0.98	1142	0.02
MEG time (s)	8.5	30709	0.04

we, however, compare the local subtraction approach to the classical subtraction approach, we see that the computation times are decreased by multiple orders of magnitude. Concretely, in the EEG case, the local subtraction approach can simulate 1000 dipoles in ca. 1 s, while in the MEG case, it can simulate 1000 dipoles in below 10 s. For *mesh_brain* in table 2.4, the MEG computation times even decrease, as, due to the higher brain resolution, larger parts of both the patch as well as the transition region are contained inside the brain compartment, where the integrals from definition 2.5.17 vanish. The corresponding computations for the classical subtraction approach, on the other hand, take multiple hours. One further difference is that the computation times for the multipolar Venant approach, as well as for the local subtraction approach, are essentially independent of the mesh size, while the computation times for the classical subtraction approach depend on the mesh size. This demonstrates that the local subtraction approach is vastly more efficient than the classical subtraction approach, which was one of the main goals of its development.

As a final remark, we want to note that the local subtraction times in the tables above are slightly too optimistic for realistic head models, as, for spherical models, a disproportionately high part of the patch and transition regions overlaps with the brain compartment, where, e.g., in the MEG case, no integral needs to be assembled. However, as these tables show the computation times for an eccentricity of 0.99, they are also not too far off. In the next subsection, we will also present local subtraction benchmarks in realistic head models.

2.8.5 Local subtraction approaches in realistic head models

In the preceding subsection, we have shown that a central benefit of the local subtraction over the classical subtraction approach is that the local subtraction approach is less dependent on the global mesh structure than the classical subtraction approach. We, in particular, saw that the accuracy of the classical subtraction approach strongly depends on the resolution of the skin compartment. We now want to illustrate this effect in a realistic head model. To this end, we took the realistic head model and its estimated dipolar generator of the somatosensory signal due to a right wrist median nerve stimulation, as described in section 2.8.1. For this model and source, we then computed the electric potential in the head, once using a classical subtraction approach, and once using a local subtraction approach. We then used the same slice as shown in figure 2.18 (b), and sampled the resulting head potentials, correction potentials, and (cut off) infinity potentials. The results for the classical subtraction approach are visualized in figure 2.30, and the results for the local subtraction approach are visualized in figure 2.31.

From figure 2.30 we can already see the high dependency of the classical subtraction approach on tissue interfaces. Especially along the axis defined by the dipole, one can see large color gradients, leading to a patchwork of different colors. Particularly noteworthy is, e.g., the red patch at the back side of the gyrus containing the source. Here, the presence of the highly conductive CSF channels the current in the head, which, as a result, is strongly different than the current in a homogeneous tissue. Accordingly, the correction potential needs to do a lot of “correcting”

Potentials involved in the classical subtraction approach

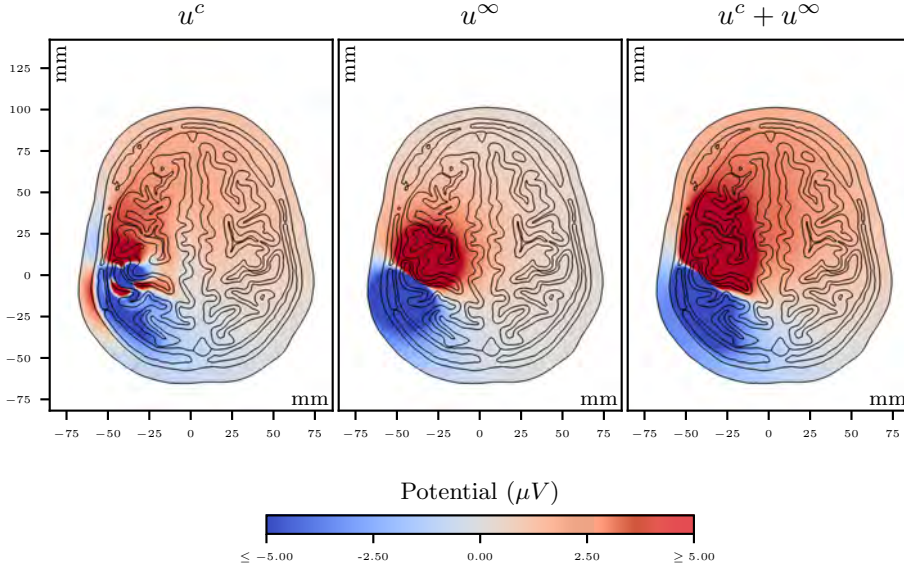


Figure 2.30: Correction potential u^c , homogeneous potential u^∞ , and head potential $u = u^c + u^\infty$, as computed by the classical subtraction approach from definition 2.5.10 for the head model and primary source as described in section 2.8.1. The potentials are evaluated by sampling the axial slice from figure 2.18 with a resolution of 0.1 mm. The black lines represent tissue boundaries and are extracted from 2.8.1 (b). This figure, and its caption, are taken in slightly adapted form from [91].

here. Furthermore, note the inhomogeneity and high gradients of the correction potential in the skull on the left-hand side of the source. As the skull is extremely low-conducting, the current tends to take the shortest way through it that is possible. This means that the current typically flows quite radially through the skull compartment. As the infinity potential is completely independent of the structure of the skull interface, this radial current needs to come from the correction potential, and we indeed see a strong radial gradient in figure 2.30 in the skull.

For the local subtraction approach in figure 2.31, we instead see a largely homogeneous correction potential that only differs from the simulated head potential in the immediate environment of the source position. These pictures thus give us a strong visual intuition for the results we saw in figures 2.26 and 2.29, namely that a good resolution in the skin compartment, and in particular at the skin-skull interface, is essential for an accurate simulation when employing the classical subtraction approach, while it is not as important for the local subtraction approach.

Additionally, when comparing the simulated head potential in figure 2.30 and figure 2.31, it is difficult to visually see any difference. Thus, even though the correction potentials look extremely different, and the two approaches take vastly different routes to perform the simulation, they arrive at an extremely similar final result. We emphasize this as we do not want to leave the impression that the classical subtraction approach is unfit to perform forward simulations. As is also shown in figure 2.29, if the mesh has a sufficiently high resolution in all critical regions, we can also expect the classical subtraction approach to produce highly accurate results. The benefit of the local subtraction approach is its increased robustness against a suboptimal mesh

Potentials involved in the local subtraction approach

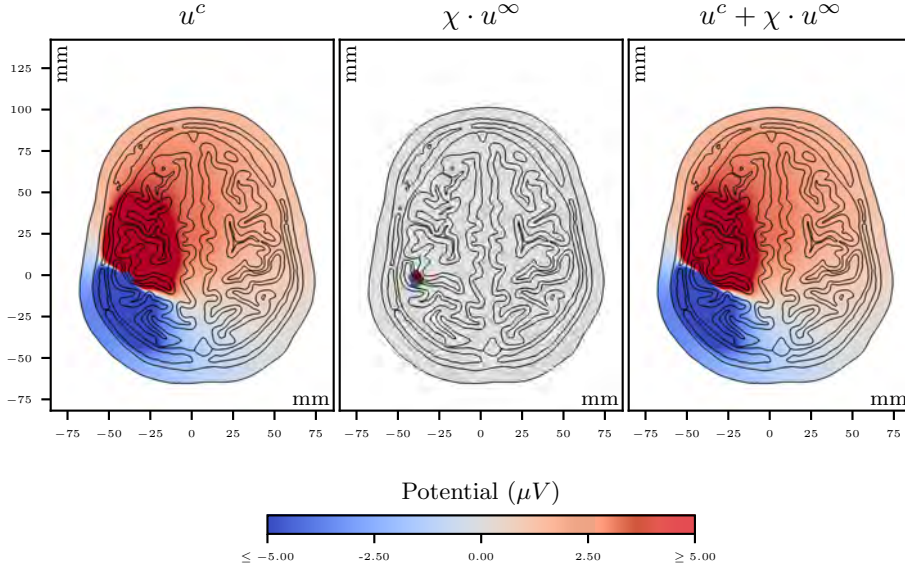


Figure 2.31: Correction potential u^c , cut off homogeneous potential $\chi \cdot u^\infty$, and head potential $u = u^c + \chi \cdot u^\infty$, as computed by the local subtraction approach from definition 2.5.12 for the head model and primary source as described in section 2.8.1. The potentials are evaluated by sampling the axial slice from figure 2.18 with a resolution of 0.1 mm. The black lines represent tissue boundaries and are extracted from 2.8.1 (b). This figure, and its caption, are taken in slightly adapted form from [91].

structure and its increased computational efficiency.

Finally, we want to present the computational effort of the local subtraction approach in a realistic head model. Following section 2.7.3, we constructed a regular source space with a resolution of 1.5 mm and rejected all positions that were not contained in the gray matter or did not have a distance of at least 0.35 mm from the gray matter compartment boundary. This led to a set of ca. 170k source positions. Out of these, we randomly selected 1000 positions. For each of these positions, we then computed both the EEG and the MEG lead field for each of the three unit directions as dipole moments. This means that we effectively solve the EEG and MEG forward problems for 3000 dipoles, in contrast to the results from tables 2.3 - 2.5, where we solved the forward problem for 1000 dipoles. The forward simulation was performed using a sensor setup from a VEP/VEF measurement experiment performed at the IBB, and consisted of 57 electrodes and 592 MEG coils. In this setup, and using an AMD Ryzen Threadripper 3960X processor with 24 cores and the local subtraction approach, the computation of the EEG lead field $L_E \in \mathbb{R}^{57 \times (3 \cdot 1000)}$ in one exemplary run took ca. 2.5 seconds, and the computation of the MEG *coil* leadfield $L_M \in \mathbb{R}^{592 \times (3 \cdot 1000)}$ took ca. 21 seconds.⁴⁰

This finishes our numerical investigations into the local subtraction approach. As a final topic in this chapter on forward modeling, we will now turn our attention to *reciprocity laws*.

⁴⁰Note that we typically apply a linear transformation to L_M to transform coils to *channels*. For the VEP/VEF experiment underlying the sensor positions, we have, e.g., 300 MEG channels.

2.9 The reciprocity laws of quasi-static bioelectromagnetism

Reciprocity principles provide a framework for exchanging the roles of source and measurement. Because of this, they form a central underpinning for both the theory and practice of studying the human brain, linking passive measurements to active stimulations. Example applications include optimizing tES montages from EEG measurements [46, 60], estimating the sensitivity distribution of electrical impedance tomography [125], or predicting what sources in the brain can be excited or inhibited using TMS from knowledge about their generated magnetic fields [81]. The original statement of reciprocity principles is attributed to Helmholtz [89]. Mathematical derivations of the concrete expressions of the reciprocity laws as they are used in quasi-static bioelectromagnetism are, e.g., given in [205] for the EEG/tES case and in [205, 140] for the MEG/TMS case. These works arrive at the reciprocity laws only after complex computations, involving intricate mathematical machinery. In this section, we will instead introduce a simple functional analytic framework that will enable us to painlessly derive a general abstract reciprocity principle. We will both state and prove this general reciprocity theorem and show how the concrete reciprocity laws of quasi-static bioelectromagnetism follow almost trivially from it.

Furthermore, we will show that the well-known transfer matrix approach, which enables a rapid forward simulation of electric and magnetic fields generated by neural currents [227, 225], is equivalent to expressing the general reciprocity principle in a coordinate system. This is a surprising connection, as the transfer matrix approach is typically introduced as a “linear algebra trick”. Additionally, this result generalizes a partial result from [225], where it is shown that the EEG adjoint approach from [205] is equivalent to an EEG transfer matrix approach coupled with a partial integration source model.

The following derivation of reciprocity principles is, to the best of our knowledge, not known in the bioelectromagnetism community. In the wider electromagnetism literature, there are papers going in a similar direction, e.g. [33]. Nevertheless, we are not aware of a work that contains the arguments presented below.

2.9.1 The general reciprocity principle

Let H denote a vector space and $a : H \times H \rightarrow \mathbb{R}$ a bilinear form. We furthermore assume that a is *symmetric*, which means that for $u, v \in H$ we have

$$a(u, v) = a(v, u).$$

Now let $\varphi : H \rightarrow \mathbb{R}$ denote a linear functional. Then φ can act as a *source* or as a *measurement*, as defined in the following.

Definition 2.9.1. We say that φ *acts as a source* if it defines a problem as follows.

$$\text{Search for } u \in H \text{ such that for all } v \in H \text{ we have } a(u, v) = \varphi(v). \quad (2.76)$$

We also call φ the *source functional* of problem (2.76).

Problems of this form are typically called *weak formulations* or *Galerkin methods*, see e.g. [25]. In the following, we will denote a solution to (2.76) as u_φ . More concretely, this means that the defining condition of $u_\varphi \in H$ is that for all $v \in H$ we have

$$a(u_\varphi, v) = \varphi(v). \quad (2.77)$$

Note that we explicitly do not assume that solutions to (2.76) are unique. E.g., in the EEG case, solutions can differ by an additive constant, representing the arbitrary choice of reference. When we write u_φ , we instead only assume that at least one solution exists and that u_φ is one of the potentially many solutions.

We say that φ acts as a measurement if we apply φ to some element $u \in H$ to compute the value $\varphi(u) \in \mathbb{R}$.

We can now formulate and prove the general reciprocity principle.

Theorem 2.9.2 (General reciprocity principle). *Let $\varphi, \psi : H \rightarrow \mathbb{R}$ denote two linear functionals. We then have*

$$\psi(u_\varphi) = \varphi(u_\psi). \quad (2.78)$$

This means that the result of applying ψ as a measurement to the solution generated by the source functional φ yields the same result as applying φ as a measurement to the solution generated by the source functional ψ . This thus enables us to exchange the roles of source and measurement.

Proof. By definition 2.76, we have

$$\psi(u_\varphi) \stackrel{\text{Definition of } u_\psi}{=} a(u_\psi, u_\varphi) \stackrel{a \text{ symmetric}}{=} a(u_\varphi, u_\psi) \stackrel{\text{Definition of } u_\varphi}{=} \varphi(u_\psi).$$

□

Note the simplicity of both the statement as well as the proof of theorem 2.9.2. Furthermore, note that each pair of functionals (φ, ψ) induces a reciprocity law. In the following, we will see that the reciprocity principles of quasi-static bioelectromagnetism arise by simply plugging the source functionals of the different modalities into theorem 2.9.2 and explicitly writing out (2.78). Before doing this, we want to highlight three more general points.

Remark 2.9.3. When deriving a FEM approach, one generally proceeds as follows. First, one chooses a suitable, typically infinite-dimensional, vector space H and shows that the function of interest can be described using a weak formulation as defined in (2.76). In FEM, the approach is then to choose a finite-dimensional subspace $H_0 \subset H$, and, instead of solving the original weak formulation

$$\text{Find } u \in H \text{ such that } a(u, v) = \varphi(v) \text{ for all } v \in H$$

on the infinite-dimensional space H , solve its restriction to the finite-dimensional subspace H_0 given by

$$\text{Find } u_0 \in H_0 \text{ such that } a(u_0, v_0) = \varphi(v_0) \text{ for all } v_0 \in H_0.$$

If we now have a pair of functionals (φ, ψ) on H , we can apply theorem 2.9.2 both to the problem on H as well as to the problem on H_0 . We thus see that a reciprocity law which is valid for a pair of solutions of the continuous problem is also valid for the corresponding pair of solutions from the FEM discretization of the problem.

Remark 2.9.4. Since we assume the form a to be bilinear, it follows that the solution of problem (2.76) depends linearly on the source functional. This means that for functionals φ_1, φ_2 and scalars $\lambda, \mu \in \mathbb{R}$ we have

$$u_{\lambda \cdot \varphi_1 + \mu \cdot \varphi_2} = \lambda \cdot u_{\varphi_1} + \mu \cdot u_{\varphi_2}.$$

Remark 2.9.5. In the phrasing of theorem 2.9.2, we assumed that a is a symmetric bilinear form on $H \times H$. This is sufficient for our purposes, but we can also formulate the theorem in a more

general context. Assume that H_1 and H_2 are two vector spaces, where it is, in general, possible that $H_1 \neq H_2$. Now, let $a : H_1 \times H_2 \rightarrow \mathbb{R}$ denote a bilinear form, which does not need to be symmetric. Then, let $\varphi : H_2 \rightarrow \mathbb{R}$ and $\psi : H_1 \rightarrow \mathbb{R}$ denote linear functionals, and let $u_\varphi \in H_1$ be a solution of

$$\text{Find } u \in H_1 \text{ such that } a(u, v) = \varphi(v) \text{ for all } v \in H_2$$

and let $u^\psi \in H_2$ be a solution of

$$\text{Find } u \in H_2 \text{ such that } a(v, u) = \psi(v) \text{ for all } v \in H_1.$$

Then, using the same argument as in theorem 2.9.2, we get

$$\psi(u_\varphi) = \varphi(u^\psi).$$

We thus get almost the same result as in theorem 2.9.2, with the difference that u^ψ above now fulfills

$$a(v, u^\psi) = \psi(v) \text{ for all } v \in H_1,$$

i.e., the test functions are inserted into the *first* argument of the bilinear form. In the setting of theorem 2.9.2, this makes no difference, as there we assumed the form to be symmetric. In the general case, this is, however, an important distinction. To give a concrete example, if we look at section 2.4, we see that, as long as a PEM model is employed, the bilinear form underlying the weak formulation is given by

$$(u, v) \mapsto a_{\text{PEM}}(u, v) := \int_{\Omega} \langle \sigma \nabla u, \nabla v \rangle dV.$$

In this thesis, we always assume that σ is symmetric, which implies that a_{PEM} is symmetric. If we however do not assume σ to be symmetric, and have a pair of functionals (φ, ψ) with corresponding reciprocal solution pair (u_φ, u^ψ) defined as above, it follows that u^ψ arises from the left hand side

$$v \mapsto a_{\text{PEM}}(v, u^\psi) = \int_{\Omega} \langle \sigma \nabla v, \nabla u^\psi \rangle dV = \int_{\Omega} \langle \sigma^\top \nabla u^\psi, \nabla v \rangle dV,$$

which implies that u^ψ solves the equation

$$\text{div}(\sigma^\top \nabla u^\psi) = -\psi \tag{2.79}$$

in a distributional sense in the interior of the volume conductor.

The main reason we present this discussion is that Guido Nolte, in his derivation of the MEG reciprocity in [140], does, at least initially, not assume the conductivity σ to be symmetric, and proceeds to derive equations in the form of (2.79). This remark thus shows that, even though [140] does not use a symmetric setting, their derivation still naturally fits inside, and is simplified by, our functional analytic framework. Furthermore, the perspective shown in this remark gives an intuitive explanation of why the so-called *lead field theorem* from [140] involves σ^\top , i.e., the transpose of the conductivity tensor, while the original equation involves σ .

2.9.2 The reciprocity of EEG and tES

In this subsection, we want to derive the reciprocity laws of the EEG and tES forward problems. As the forward problems can be phrased using a PEM as well as a CEM approach, we also get PEM and CEM reciprocity laws. We can, in fact, use the reciprocity laws to deduce the

convergence of the EEG CEM solution to the EEG PEM solution from the corresponding result for the tES case. We will first start with the PEM case, then continue with the CEM case, and then present the proof that the EEG CEM solution converges to the EEG PEM solution.

In section 2.5, we have discussed that the EEG forward problem can be reduced to simulating the electric potential due to a current dipole. A current dipole is given by a *position* $x_0 \in \mathbb{R}^3$ and a *moment* $M \in \mathbb{R}^3$. In the derivations leading up to definition 2.5.1, we saw that the weak formulation of the PEM EEG forward problem for a single current dipole is then given by

$$\text{Find } u \in H \text{ such that } \int_{\Omega} \langle \sigma \nabla u, \nabla v \rangle dV = -\operatorname{div}(M \cdot \delta_{x_0})(v) \left(= \langle M, \nabla v(x_0) \rangle \right) \text{ for all } v \in H. \quad (2.80)$$

Here, H is some suitable function space, Ω is the head domain, σ is the conductivity tensor, and δ_{x_0} is the Dirac distribution at x_0 .

We thus see that the bilinear form for the PEM EEG forward problem is given by

$$a_{\text{PEM}}(u, v) = \int_{\Omega} \langle \sigma \nabla u, \nabla v \rangle dV.$$

Since the conductivity tensor σ is symmetric, we see that a is also symmetric. Furthermore, we see that the source functional of a current dipole is given by

$$\varphi^{\text{dip}}(v) := \langle M, \nabla v(x_0) \rangle. \quad (2.81)$$

Let $u^{\text{dip}} := u_{\varphi^{\text{dip}}}$ denote a solution of (2.80). In EEG, one measures the potential differences between electrodes. If we have an electrode at position $r_1 \in \partial\Omega$ and a reference electrode at position $r_0 \in \partial\Omega$, the EEG measurement is thus given by

$$u^{\text{dip}}(r_1) - u^{\text{dip}}(r_0) = (\delta_{r_1} - \delta_{r_0})(u^{\text{dip}}).$$

In EEG, the measurement functional is hence given by

$$\psi^{\text{EEG}}(v) := \int_{\partial\Omega} (\delta_{r_1} - \delta_{r_0}) \cdot v dS = v(r_1) - v(r_0).$$

In the PEM EEG forward problem, we are thus interested in evaluating the measurement functional ψ^{EEG} on the solution corresponding to the source functional φ^{dip} with respect to the bilinear form a_{PEM} .

In tES, we are interested in simulating the electric potential (resp. the electric field) in the head which arises due to an electric current applied at the skin. If we use a PEM model, we saw in definition 2.4.1 that if we perform an electric stimulation with an anode at position $r_1 \in \partial\Omega$ and a cathode at position $r_0 \in \partial\Omega$ with a current amplitude of I , the corresponding weak formulation is given by

$$\text{Find } u \in H \text{ such that } \int_{\Omega} \langle \sigma \nabla u, \nabla v \rangle dV = I \cdot v(r_1) - I \cdot v(r_0) \text{ for all } v \in H, \quad (2.82)$$

where again H is a suitable function space, Ω is the head domain, σ is the conductivity tensor, and δ_{r_i} is the Dirac distribution at position r_i .

We thus see that the bilinear form for the tES forward problem is again given by

$$a_{\text{PEM}}(u, v) = \int_{\Omega} \langle \sigma \nabla u, \nabla v \rangle dV,$$

and that the source functional for a tES stimulation with an anode at r_1 and a cathode at r_0 with a stimulation amplitude of I is given by

$$\varphi^{\text{tES}}(v) := I \cdot (v(r_1) - v(r_0)).$$

Let $u^{\text{tES}} := u_{\varphi^{\text{tES}}}$ denote a solution of (2.82).

Note that, until now, we have only recapitulated the weak formulations of the PEM EEG and tES forward problems and introduced the necessary terminology. We will now derive the PEM EEG/tES reciprocity law.

Theorem 2.9.6 (PEM EEG/tES reciprocity law). *Let Ω denote the head domain. Let u^{dip} denote a PEM EEG forward solution due to a dipole at position $x_0 \in \Omega$ with moment $M \in \mathbb{R}^3$. Let u^{tES} denote a PEM tES forward solution due to a stimulation with amplitude I using an anode at position $r_1 \in \partial\Omega$ and a cathode at position $r_0 \in \partial\Omega$. We then have*

$$I \cdot (u^{\text{dip}}(r_1) - u^{\text{dip}}(r_0)) = \langle M, \nabla u^{\text{tES}}(x_0) \rangle. \quad (2.83)$$

Proof. First, note that the EEG measurement functional and the tES source functional are related by

$$\varphi^{\text{tES}} = I \cdot \psi^{\text{EEG}}.$$

As shown above, in EEG, we apply the measurement functional ψ^{EEG} to the solution u^{dip} generated by the source functional φ^{dip} . We can thus apply the general reciprocity principle 2.9.2, and get

$$\begin{aligned} u^{\text{dip}}(r_1) - u^{\text{dip}}(r_0) &= \psi^{\text{EEG}}(u_{\varphi^{\text{dip}}}) \stackrel{\text{Theorem 2.9.2}}{=} \varphi^{\text{dip}}(u_{\psi^{\text{EEG}}}) \stackrel{\text{Remark 2.9.4}}{=} \frac{1}{I} \cdot \varphi^{\text{dip}}(u_{\varphi^{\text{tES}}}) \\ &\stackrel{\text{Definition (2.81) of } \varphi^{\text{dip}}}{=} \frac{\langle M, \nabla u^{\text{tES}}(x_0) \rangle}{I}. \end{aligned}$$

□

We want to make a few comments.

Remark 2.9.7. By the argument of Remark 2.9.3, the reciprocity law (2.83) is not only valid for the continuous EEG and tES solutions but also for their FEM approximations. More concretely, if u_h^{tES} denotes a FEM approximation of the tES forward problem with an amplitude of $I = 1$, and u_h^{dip} denotes the solution of applying a FEM approach directly to (2.80), we have

$$u_h^{\text{dip}}(r_1) - u_h^{\text{dip}}(r_0) = \langle M, \nabla u_h^{\text{tES}}(x_0) \rangle.$$

Now note that the computation of u_h^{tES} does not depend on any dipole position or moment. We can thus compute the FEM tES solution once, and then, for *arbitrary dipoles*, compute potential differences of the generated electric potentials by performing a simple computation on the FEM tES solution, without needing to explicitly compute the FEM solution for the EEG forward problem. From a top-level perspective, this achieves the same effect as using a transfer matrix

approach [238]. Later on, we will see that this is not only similar to a transfer matrix approach but is *the exact same* as using a transfer matrix approach.

Note that, in numerical practice, we, in general, do not directly apply a FEM approach to (2.80). In the EEG FEM community, doing this is called the *partial integration approach* (see definition 2.5.1) and has some numerical shortcomings. One instead uses a functional which interacts with the FEM function space in a better way [12]. We refer to the preceding subsections of the thesis for an in-depth discussion of this question.

Remark 2.9.8. We want to note that the argument presented above has a technical gap in it if one tries to apply it to the continuous case. Concretely, we did not specify how exactly to choose the trial and test spaces in the continuous formulation. To arrive at a completely rigorous argument, we believe that one needs to use the reciprocity formulation as it is outlined in remark 2.9.5 and show that one can define spaces U and V such that U contains the dipole solutions, V contains the tES solutions, the dipole solutions can be tested against functions from V , and the tES solutions can be tested against functions from U . As the dipole solution is singular in the interior, but regular on the boundary, and the tES solution is singular on the boundary, but regular in the interior, we believe that such a construction is possible. We will, however, not try to give such a construction in this thesis. Furthermore, we want to emphasize that this is not a problem specific to our functional analytic framework, but applies to every single derivation of the reciprocity laws that we are aware of.

We now continue with the CEM reciprocity law. We, again, first quickly repeat the corresponding weak formulations. We use the notation from definition 2.4.2, i.e. we have L electrodes with disjoint contact interfaces $e_1, \dots, e_L \subset \partial\Omega$ and contact impedances $0 < z_1, \dots, z_L \in \mathbb{R}$. Furthermore, let $\mathcal{H} := H^1(\Omega) \oplus \mathbb{R}^L$, and define a bilinear form $a_{\text{CEM}} : \mathcal{H} \times \mathcal{H} \rightarrow \mathbb{R}$ via

$$a_{\text{CEM}}((u, U), (v, V)) := \int_{\Omega} \langle \sigma \nabla u, \nabla v \rangle dV + \sum_{l=1}^L \frac{1}{z_l} \cdot \int_{e_l} (u - U_l)(v - V_l) dS$$

Following definition 2.5.1, the dipole functional for a current dipole at position x_0 with moment M is again given by

$$\varphi^{\text{dip}}(v, V) = \langle M, \nabla v(x_0) \rangle.$$

But in (slight) contrast to the PEM case, the measurement functional is now given by the difference between the potential values in the electrodes, i.e., if we are interested in the potential difference between the i -th and j -th electrode, we measure

$$\psi^{\text{EEG}}(u, U) = U_i - U_j.$$

Written succinctly, the CEM EEG forward problem consists of applying the measurement functional ψ^{EEG} to the solution generated by the source functional φ^{dip} with respect to the bilinear form a_{CEM} .

In definition 2.4.2, we have furthermore derived that if in such a CEM model we want to use the i -th electrode as an anode and the j -th electrode as a cathode and insert a current of amplitude I , the corresponding electric potential $(u^{\text{tES}}, U^{\text{tES}})$ is given by the weak formulation

$$a_{\text{CEM}}((u^{\text{tES}}, U^{\text{tES}}), (v, V)) = I \cdot V_i - I \cdot V_j,$$

i.e. $(u^{\text{tES}}, U^{\text{tES}})$ arises from the source functional

$$\varphi^{\text{tES}}(v, V) = I \cdot (V_i - V_j).$$

with respect to the bilinear form a_{CEM} . We thus again see that the tES source functional is simply a multiple of the EEG measurement functional, i.e. $\varphi^{\text{tES}} = I \cdot \psi^{\text{EEG}}$, which again implies the corresponding reciprocity principle.

Theorem 2.9.9 (CEM EEG/tES reciprocity law). *Let Ω denote the head domain. Assume that we employ $L \in \mathbb{N}$ electrodes with disjoint contact interfaces $e_1, \dots, e_L \subset \partial\Omega$ and contact impedances $0 < z_1, \dots, z_L$. Let $(u^{\text{dip}}, U^{\text{dip}})$ denote a CEM EEG forward solution due to a dipole at position $x_0 \in \Omega$ with moment $M \in \mathbb{R}^3$. Let $(u^{\text{tES}}, U^{\text{tES}})$ denote a CEM tES forward solution due to a stimulation with amplitude I using the i -th electrode as an anode and the j -th electrode as a cathode. We then have*

$$I \cdot (U_i^{\text{dip}} - U_j^{\text{dip}}) = \langle M, \nabla u^{\text{tES}}(x_0) \rangle. \quad (2.84)$$

Proof. Just as in the proof of theorem 2.9.6, we apply the general reciprocity principle 2.9.2 to get

$$\begin{aligned} U_i^{\text{dip}} - U_j^{\text{dip}} &= \psi^{\text{EEG}}((u^{\text{dip}}, U^{\text{dip}})) \stackrel{\text{Theorem 2.9.2}}{=} \varphi^{\text{dip}}((u_{\psi^{\text{EEG}}}, U_{\varphi^{\text{dip}}})) \\ &\stackrel{\text{Remark 2.9.4}}{=} \frac{1}{I} \cdot \varphi^{\text{dip}}((u_{\varphi^{\text{tDCS}}}, U_{\varphi^{\text{tDCS}}})) \stackrel{\text{Definition of } \varphi^{\text{dip}}}{=} \frac{\langle M, \nabla u^{\text{tDCS}}(x_0) \rangle}{I}. \end{aligned}$$

□

To illustrate how much the theory of reciprocity laws is simplified by our functional analytic framework, we invite the reader to try and derive theorem 2.9.9 directly from the differential formulations 2.1.2 and 2.2.2.

As a final topic in this subsection, we now want to prove the convergence of the CEM EEG solution to the PEM EEG solution. We want to remind the reader of definition 2.4.7, where it was specified what we mean by “shrinking” electrodes. Furthermore, we remind the reader of theorem 2.4.8, where we cited a result from [84] showing the convergence of the CEM tES solution towards the PEM tES solutions. We will now use reciprocity laws to extend the result to the EEG problem.

As a further preparation, remember that both in the CEM case, as well as in the PEM case, the solution of the EEG forward problem is only determined up to an additive constant. If we denote $\mathbf{1} = (1, \dots, 1) \in \mathbb{R}^L$, and we gather the values of the potential at a selection of L positions in a vector $U = (u_1, \dots, u_L) \in \mathbb{R}^L$, this means that while U is not uniquely determined, its equivalence class in the quotient space $\mathbb{R}^L / \mathbb{R} \cdot \mathbf{1}$ is uniquely determined.

Theorem 2.9.10. *Let $x_1, \dots, x_L \in \partial\Omega$ denote point electrode positions and let $e_1^h, \dots, e_L^h \subset \partial\Omega$ denote electrode interfaces parametrized by $h \in (0, h_0)$. Furthermore, let $x_0 \in \Omega$ denote a dipole position and $M \in \mathbb{R}^3$ a dipole moment. Let $(u_{\text{CEM}}^h, U_{\text{CEM}}^h)$ denote a corresponding CEM EEG solution and u_{PEM} a corresponding PEM EEG solution. Define the PEM measurement vector $U_{\text{PEM}} := (u_{\text{PEM}}(x_1), \dots, u_{\text{PEM}}(x_L)) \in \mathbb{R}^L$. We then have the following.*

- *If the electrodes shrink towards x_1, \dots, x_L , there is a constant $C > 0$, independent of h and M , such that for all $0 < h < h_0$ we have*

$$\|U_{\text{CEM}}^h - U_{\text{PEM}}\|_{\mathbb{R}^L / \mathbb{R} \cdot \mathbf{1}} \leq C \cdot h \cdot \|M\|_2.$$

- If the electrodes shrink nicely towards x_1, \dots, x_L , there is a constant $C > 0$, independent of h and M , such that for all $0 < h < h_0$ we have

$$\|U_{CEM}^h - U_{PEM}\|_{\mathbb{R}^L / \mathbb{R} \cdot 1} \leq C \cdot h^2 \cdot \|M\|_2.$$

We have followed the convention of [84] and expressed the convergence results using the quotient norm on the space $\mathbb{R}^L / \mathbb{R} \cdot 1$. In a less technical, but equivalent, way, you can also say that the common average referenced CEM potentials converge entrywise towards the common average referenced PEM potentials, with a rate of either h or h^2 .

Proof. By definition of the quotient norm, we have

$$\begin{aligned} \|U_{CEM}^h - U_{PEM}\|_{\mathbb{R}^L / \mathbb{R} \cdot 1} &\leq \|U_{CEM}^h - U_{PEM} - ((U_{CEM}^h)_1 - (U_{PEM})_1) \cdot 1\|_2 \\ &\leq C \cdot \max_{2 \leq i \leq L} |((U_{CEM}^h)_i - (U_{CEM}^h)_1) - ((U_{PEM})_i - U_{PEM})_1|. \end{aligned}$$

Now, let $u_{i,j}^{\text{tES},h}$ denote the CEM tES solution for a current injection of $I = 1$, using the i -th electrode as an anode and the j -th electrode as a cathode. Similarly, let $u_{i,j}^{\text{tES},\text{PEM}}$ denote the PEM tES solution for a current injection of $I = 1$ using the point electrode position x_i as an anode and the point electrode position x_j as a cathode. Then, we have by the PEM and CEM reciprocity laws in theorems 2.9.6 and 2.9.9 that for $2 \leq i \leq L$

$$\begin{aligned} (U_{CEM}^h)_i - (U_{CEM}^h)_1 &= \langle M, \nabla u_{i,1}^{\text{tES},h}(x_0) \rangle \\ (U_{PEM})_i - U_{PEM} &= \langle M, \nabla u_{i,1}^{\text{tES},\text{PEM}}(x_0) \rangle \end{aligned}$$

This yields

$$\begin{aligned} |((U_{CEM}^h)_i - (U_{CEM}^h)_1) - ((U_{PEM})_i - U_{PEM})_1| &= |\langle M, \nabla u_{i,1}^{\text{tES},h}(x_0) - \nabla u_{i,1}^{\text{tES},\text{PEM}}(x_0) \rangle| \\ &\leq \|M\|_2 \cdot \|\nabla (u_{i,1}^{\text{tES},h} - u_{i,1}^{\text{tES},\text{PEM}})(x_0)\|_2 \\ &\leq \|M\|_2 \cdot \|u_{i,1}^{\text{tES},h} - u_{i,1}^{\text{tES},\text{PEM}} - c\|_{C^1(B_\epsilon(x_0))}, \end{aligned}$$

where $c \in \mathbb{R}$ is arbitrary and $\epsilon > 0$ is small enough such that on the ball of radius ϵ around x_0 the conductivity σ is smooth. Now, note that on $B_\epsilon(x_0)$ we have

$$\operatorname{div} \sigma \nabla (u_{i,1}^{\text{tES},h} - u_{i,1}^{\text{tES},\text{PEM}} - c) = 0$$

in a weak sense, and hence elliptic regularity (see e.g. [57], 6.3.1, Theorem 3) implies that we have $u_{i,1}^{\text{tES},h} - u_{i,1}^{\text{tES},\text{PEM}} - c \in C^\infty(B_\epsilon(x_0))$. After potentially shrinking ϵ , we thus have $u_{i,1}^{\text{tES},h} - u_{i,1}^{\text{tES},\text{PEM}} - c \in H^r(B_\epsilon(x_0))$ for an arbitrary $r \in \mathbb{R}$. But by general Sobolev inequalities (see e.g. [57], 5.6.3, Theorem 6) we then get

$$\|u_{i,1}^{\text{tES},h} - u_{i,1}^{\text{tES},\text{PEM}} - c\|_{C^1(B_\epsilon(x_0))} \leq C \cdot \|u_{i,1}^{\text{tES},h} - u_{i,1}^{\text{tES},\text{PEM}} - c\|_{H^r(B_\epsilon(x_0))}$$

for a sufficiently large r . Since $c \in \mathbb{R}$ was arbitrary, this yields

$$\|U_{CEM}^h - U_{PEM}\|_{\mathbb{R}^L / \mathbb{R} \cdot 1} \leq \|M\|_2 \cdot C \cdot \max_{2 \leq i \leq L} \|u_{i,1}^{\text{tES},h} - u_{i,1}^{\text{tES},\text{PEM}}\|_{H^r(B_\epsilon(x_0)) / \mathbb{R} \cdot 1}$$

Now, in [66], Lemma 9, it is discussed how to extend Lemma A.1 from [85] to show that, if $\epsilon' > \epsilon$ is small enough such that σ is smooth on $B_{\epsilon'}(x_0)$, there is a constant $C > 0$ such that for all $v \in H^1(B_{\epsilon'}(x_0))$ with $\operatorname{div}(\sigma \nabla v) = 0$ we have

$$\|v\|_{H^r(B_{\epsilon}(x_0))_{/\mathbb{R},1}} \leq C \cdot \|v\|_{H^1(B_{\epsilon'}(x_0))_{/\mathbb{R},1}},$$

which, in combination with the preceding computation, yields

$$\|U_{\text{CEM}}^h - U_{\text{PEM}}\|_{\mathbb{R}^L_{/\mathbb{R},1}} \leq \|M\|_2 \cdot C \cdot \max_{2 \leq i \leq L} \|u_{i,1}^{\text{tES},h} - u_{i,1}^{\text{tES,PEM}}\|_{H^1(B_{\epsilon'}(x_0))_{/\mathbb{R},1}}.$$

But now, the statement of the theorem follows from theorem 2.4.8. \square

The argument above is, to the best of our knowledge, the first proof of the convergence of the CEM EEG solution towards the PEM EEG solution in the continuous case. The only other argument going in a similar direction, that we are aware of, is given in [160], where the authors show that, *in the discrete case*, if one first takes the limit of the impedances z_1, \dots, z_L going to infinity, and then shrinks the CEM electrode interfaces towards points, the final result will be the PEM solution. The step of first letting the impedances approach infinity is, in our opinion, not motivated, and, given the results from theorems 2.4.8 and 2.9.10, is also not necessary. In fact, looking at the argument in [160], it is not difficult to adapt it in such a way that one can directly derive the convergence of the discrete EEG CEM solution towards the discrete EEG PEM solution for a fixed impedance setup, without needing to first let them go to infinity. As this thesis is already getting quite long, we will leave this as an exercise for the reader.

2.9.3 The reciprocity of MEG and TMS

We will derive two MEG/TMS reciprocity laws, one for the magnetic field itself and one for the magnetic flux. Similar to the EEG/tES case, this will consist of identifying MEG and TMS source and measurement functionals, noticing that the MEG measurement functional is essentially the TMS source functional, and applying theorem 2.9.2. In this subsection, we will always assume a PEM model.

The MEG forward problem

In definition 2.1.3, we have derived that the magnetic field B at position z in direction n due to a neural activity j^P is given by

$$\begin{aligned} \langle B(z), n \rangle &= \left\langle \frac{\mu_0}{4\pi} \cdot \int_{\Omega} j(y) \times \frac{z-y}{\|z-y\|^3} dV(y), n \right\rangle \\ &= \underbrace{\left\langle \frac{\mu_0}{4\pi} \cdot \int_{\Omega} j^P(y) \times \frac{z-y}{\|z-y\|^3} dV(y), n \right\rangle}_{=: B^P(z)} + \underbrace{\left\langle -\frac{\mu_0}{4\pi} \cdot \int_{\Omega} \sigma \nabla u(y) \times \frac{z-y}{\|z-y\|^3} dV(y), n \right\rangle}_{=: B^S(z)} \\ &= \langle B^P(z), n \rangle + \langle B^S(z), n \rangle. \end{aligned}$$

This introduces a splitting of the magnetic flux density B into a primary field B^P and a secondary field B^S .

As before, the forward problem can be reduced to simulating the magnetic field due to a single current dipole, given by its position and moment. Let $x_0 \in \Omega$ denote a dipole position and

$M \in \mathbb{R}^3$ a dipole moment. Then, the neural activity of this point dipole is given by $j^P = M \cdot \delta_{x_0}$. We thus get for the primary magnetic flux density of the dipole that

$$\begin{aligned} B^P(z) &= \frac{\mu_0}{4\pi} \cdot \int_{\Omega} (M \cdot \delta_{x_0}(y)) \times \frac{z - y}{\|z - y\|^3} dV(y) \\ &= \frac{\mu_0}{4\pi} \cdot M \times \frac{z - x_0}{\|z - x_0\|^3}. \end{aligned}$$

The primary B-field of a current dipole thus has a simple analytic expression, and the mathematically difficult-to-handle part is given by the secondary B-field. Now, just as in the EEG case in equation (2.80), the electric potential u^{dip} generated by the dipole is given by the problem

$$\text{Find } u \in H \text{ such that } \int_{\Omega} \langle \sigma \nabla u, \nabla v \rangle dV = \langle M, \nabla v(x_0) \rangle \text{ for all } v \in H,$$

for a suitable function space H . If we now define a measurement functional

$$\psi^{\text{MEG}}(v) = \langle -\frac{\mu_0}{4\pi} \cdot \int \sigma \nabla v \times \frac{z - y}{\|z - y\|^3} dV(y), n \rangle, \quad (2.85)$$

we see that the secondary B-field contribution to the MEG measurement of the current dipole is given by applying the measurement functional ψ^{MEG} to the solution generated by the source functional φ^{dip} , as defined in equation (2.81).

The TMS forward problem

In definition 2.4.10, we derived that the TMS forward problem consists of computing the electric field

$$E = -\frac{\partial A}{\partial t} - \nabla u, \quad (2.86)$$

generated by a magnetic vector potential A . Furthermore, we saw that the potential u in the equation above is described by the weak formulation

$$\text{Find } u \in H \text{ such that } \int_{\Omega} \langle \sigma \nabla u, \nabla v \rangle dV = \int_{\Omega} \langle -\sigma \frac{\partial A}{\partial t}, \nabla v \rangle dV \text{ for all } v \in H. \quad (2.87)$$

In the following, we define $E^P := -\frac{\partial A}{\partial t}$ as the *primary electric field*, and $E^S := -\nabla u$ as the *secondary electric field*, and we hence have a decomposition of the TMS electric field into the sum of a primary and a secondary field.

In the discussion preceding definition 2.2.3, we saw that the magnetic vector potential can be written as $A(x, t) = f(t) \cdot A^s(x)$, where A^s is the *spatial component* and f is the *time course* of the magnetic vector potential. Furthermore, as discussed immediately after definition 2.2.3, a common strategy is to model the vector potential A^s as arising from a linear combination of magnetic dipoles [202, 47, 123]. A magnetic dipole is given by a position $z \in \mathbb{R}^3$ and a moment $n \in \mathbb{R}^3$, and, as previously discussed in (2.17), induces a spatial vector potential

$$A^S(y) = \frac{\mu_0}{4\pi} \cdot n \times \frac{y - z}{\|y - z\|^3}. \quad (2.88)$$

From equation (2.87), we thus get that the TMS forward problem for a magnetic dipole with position $z \in \mathbb{R}^3$, moment $n \in \mathbb{R}^3$, and time course $f(t)$ is given by the weak formulation

Find $u \in H$ such that for all $v \in H$ we have

$$\int_{\Omega} \langle \sigma \nabla u, \nabla v \rangle dV = -f'(t) \cdot \frac{\mu_0}{4\pi} \cdot \int_{\Omega} \langle \sigma \left(n \times \frac{y-z}{\|y-z\|^3} \right), \nabla v \rangle dV \quad (2.89)$$

where again H is a suitable function space, Ω is the head domain, and σ is the conductivity tensor. If we thus define a source functional

$$\varphi^{\text{TMS}}(v) = -f'(t) \cdot \frac{\mu_0}{4\pi} \cdot \int_{\Omega} \langle \sigma \left(n \times \frac{y-z}{\|y-z\|^3} \right), \nabla v \rangle dV, \quad (2.90)$$

we see that, for a magnetic dipole, the potential giving the secondary field contribution to the electric field in (2.86) arises from the source functional φ^{TMS} and the PEM bilinear form $a_{\text{PEM}}(u, v) = \int_{\Omega} \langle \sigma \nabla u, \nabla v \rangle dV$.

Field version of the MEG/TMS reciprocity

After having recapitulated the MEG and TMS forward problems, we can now state and prove the MEG/TMS reciprocity theorem for the magnetic field.

Theorem 2.9.11 (MEG/TMS reciprocity law, field version). *Let Ω denote the head domain. Let B^P, B^S , and B denote the primary, secondary, and total magnetic flux densities due to a current dipole at position $x_0 \in \Omega$ with moment $M \in \mathbb{R}^3$, as defined by the MEG forward problem. Furthermore, let E^P, E^S , and E denote the primary, secondary, and total electric field due to a magnetic dipole at position $z \in \mathbb{R}^3$ with moment $n \in \mathbb{R}^3$ and timecourse $f(t)$, as defined by the TMS forward problem. We then have*

$$-f'(t) \cdot \langle B(z), n \rangle = \langle E(x_0, t), M \rangle \quad (2.91)$$

$$-f'(t) \cdot \langle B^P(z), n \rangle = \langle E^P(x_0, t), M \rangle \quad (2.92)$$

$$-f'(t) \cdot \langle B^S(z), n \rangle = \langle E^S(x_0, t), M \rangle \quad (2.93)$$

Note that the reciprocity relation is valid for the primary, the secondary, and the total fields.

Proof. Let $a, b, c \in \mathbb{R}^3$. It is well known, and otherwise easy to compute, that then $a \times b = -b \times a$, and $\langle a \times b, c \rangle = \langle b \times c, a \rangle$. Since the conductivity tensor σ is symmetric, this implies

$$\begin{aligned} \langle \sigma \left(n \times \frac{y-z}{\|y-z\|^3} \right), \nabla v \rangle &= \left\langle \left(n \times \frac{y-z}{\|y-z\|^3} \right), \sigma \nabla v \right\rangle = \left\langle \frac{y-z}{\|y-z\|^3} \times (\sigma \nabla v), n \right\rangle \\ &= \langle (\sigma \nabla v) \times \frac{z-y}{\|z-y\|^3}, n \rangle. \end{aligned}$$

This shows that the TMS source functional φ^{TMS} , as defined in (2.90), and the MEG measurement functional ψ^{MEG} for a coil at position z with orientation n , as defined in (2.85), are related via $\varphi^{\text{TMS}} = f'(t) \cdot \psi^{\text{MEG}}$. Now let $u^{\text{dip}}(:= u_{\varphi^{\text{dip}}})$ denote the solution of the EEG forward problem for the current dipole at position x_0 with moment M , and let $u^{\text{TMS}} := u_{\varphi^{\text{TMS}}}$ denote the solution of the TMS forward problem at timepoint t for a magnetic dipole at position z with moment n and timecourse $f(t)$. As described above, the secondary B-field contribution of the MEG measurement is given by applying the measurement functional ψ^{MEG} to u^{dip} . By the

general reciprocity theorem 2.9.2, we now have

$$\begin{aligned} -f'(t) \cdot \langle B^S(z), n \rangle &= -f'(t) \cdot \psi^{\text{MEG}}(u_{\varphi^{\text{dip}}}) \stackrel{\text{Theorem 2.9.2}}{=} -f'(t) \cdot \varphi^{\text{dip}}(u_{\psi^{\text{MEG}}}) \\ &\stackrel{\text{Rem. 2.9.4}}{=} -\varphi^{\text{dip}}(u_{f'(t) \cdot \psi^{\text{MEG}}}) = -\varphi^{\text{dip}}(u^{\text{TMS}}) \\ &\stackrel{\text{Eq. (2.81)}}{=} -\langle M, \nabla u^{\text{TMS}}(x_0) \rangle = \langle E^S(x_0, t), M \rangle. \end{aligned}$$

We have thus derived the reciprocity law (2.93) for the secondary fields. For the primary fields, we can similarly compute

$$\begin{aligned} -f'(t) \cdot \langle B^P(z), n \rangle &= -f'(t) \cdot \frac{\mu_0}{4\pi} \cdot \langle M \times \frac{z - x_0}{\|z - x_0\|^3}, n \rangle \\ &= \langle -f'(t) \cdot \left(\frac{\mu_0}{4\pi} \cdot n \times \frac{x_0 - z}{\|x_0 - z\|^3} \right), M \rangle \stackrel{\text{Eq. (2.88)}}{=} \langle -\frac{\partial A}{\partial t}(x_0, t), M \rangle \\ &= \langle E^P(x_0, t), M \rangle. \end{aligned}$$

This shows the reciprocity law (2.92) for the primary fields. This in turn implies

$$\begin{aligned} -f'(t) \cdot \langle B(z), n \rangle &= -f'(t) \cdot (\langle B^P(z), n \rangle + \langle B^S(z), n \rangle) = \langle E^P(x_0, t), M \rangle + \langle E^S(x_0, t), M \rangle \\ &= \langle E(x_0, t), M \rangle, \end{aligned}$$

which shows the reciprocity law (2.91) for the total fields. \square

Again, note how, after simply stating the weak formulations of the MEG/TMS forward problems, the reciprocity relations in theorem 2.9.11 are an almost trivial consequence of the general reciprocity theorem 2.9.2, coupled with a few straightforward algebraic manipulations.

Additionally, note that, just as in the EEG/tES case, the reciprocity law is not only valid for the continuous fields but also for their FEM approximations.

Flux version of the MEG/TMS reciprocity

In SQUID MEG, we do not measure the magnetic field B directly, but instead measure a value derived from the magnetic flux through a pickup coil [83]. As a final result in this section, we want to show that theorem 2.9.2 also immediately implies a reciprocity law for the magnetic flux.

Let C denote the boundary of a coil enclosing a surface F with surface normal η . When using this coil for measurements, we are interested in the magnetic flux through the coil, which is defined as

$$\Phi_B(C) := \int_F \langle B, \eta \rangle dS. \quad (2.94)$$

By inserting the splitting $B = B^P + B^S$ into this equation, we get a splitting $\Phi_B(C) = \Phi_B^P(C) + \Phi_B^S(C)$ of the magnetic flux into the sum of a *primary flux* $\Phi_B^P(C) := \int_F \langle B^P, \eta \rangle dS$ and a secondary flux $\Phi_B^S(C) := \int_F \langle B^S, \eta \rangle dS$.

On the other hand, we could use the coil for TMS stimulation by supplying it with a time-varying current. Then, as described in the TMS section above, it induces a vector potential $A(y, t) = f(t) \cdot A^S(y)$. The spatial component A^S of this vector potential is now given by

$$A^S(y) = \frac{\mu_0}{4\pi} \cdot \int_C \frac{\tau}{\|y - z\|} ds(z), \quad (2.95)$$

where τ denotes the tangential vector of the curve C , which follows from equation (2.8) (or [169], in the case you want a paper citation).

We can now formulate the flux version of the MEG/TMS reciprocity law.

Theorem 2.9.12 (MEG/TMS reciprocity law, flux version). *Let Ω denote the head volume. Let C denote the boundary curve of a coil. Let $\Phi_B^P(C)$, $\Phi_B^S(C)$, and $\Phi_B(C)$ denote the primary, secondary, and total magnetic flux through the coil C due to a current dipole at position $x_0 \in \Omega$ with moment $M \in \mathbb{R}^3$. Furthermore, let E^P , E^S , and E denote the primary, secondary, and total electric field due to a TMS stimulation using the coil C with a time course f . We then have*

$$-f'(t) \cdot \Phi_B(C) = \langle E(x_0, t), M \rangle \quad (2.96)$$

$$-f'(t) \cdot \Phi_B^P(C) = \langle E^P(x_0, t), M \rangle \quad (2.97)$$

$$-f'(t) \cdot \Phi_B^S(C) = \langle E^S(x_0, t), M \rangle \quad (2.98)$$

Proof. We again start with the secondary flux/field. Let F denote a surface enclosed by the coil C . Let u^{dip} denote the solution of the EEG forward problem for a current dipole at position x_0 with moment M . By writing out the definition of the secondary field B^S , we have

$$\Phi_B^S(C) = -\frac{\mu_0}{4\pi} \int_F \int_\Omega \langle \sigma \nabla u^{\text{dip}}(y) \times \frac{z-y}{\|z-y\|^3}, \eta(z) \rangle dV(y) dS(z),$$

and we see that if we define

$$\psi_C^{\text{MEG}}(v) := -\frac{\mu_0}{4\pi} \int_F \int_\Omega \langle \sigma \nabla v(y) \times \frac{z-y}{\|z-y\|^3}, \eta(z) \rangle dV(y) dS(z),$$

the secondary magnetic flux $\Phi_B^S(C)$ is given by applying the measurement functional ψ_C^{MEG} to the solution generated by the source functional φ^{dip} . This enables us to apply theorem 2.9.2 again, and we get

$$\Phi_B^S(C) = \psi_C^{\text{MEG}}(u_{\varphi^{\text{dip}}}) \stackrel{\text{Theorem 2.9.2}}{=} \varphi^{\text{dip}}(u_{\psi_C^{\text{MEG}}}). \quad (2.99)$$

Now, by equation (2.87), the weak formulation for the TMS forward problem for the vector potential

$$A(y, t) = f(t) \cdot A^S(y) = f(t) \cdot \frac{\mu_0}{4\pi} \cdot \int_C \frac{\tau}{\|y-z\|} ds(z)$$

is given by the source functional

$$\begin{aligned} \varphi_C^{\text{TMS}}(v) &:= \int_\Omega \left\langle -\sigma \left(f'(t) \cdot \frac{\mu_0}{4\pi} \cdot \int_C \frac{\tau}{\|y-z\|} ds(z) \right), \nabla v(y) \right\rangle dV(y) \\ &= -f'(t) \cdot \frac{\mu_0}{4\pi} \cdot \int_\Omega \left\langle \sigma \int_C \frac{\tau}{\|y-z\|} ds(z), \nabla v(y) \right\rangle dV(y). \end{aligned}$$

Now, it is well known, and otherwise easily derived from a component-wise application of Stokes' theorem, that for a scalar field f we have

$$\int_F \eta \times \nabla f dS = \int_C f \cdot \tau ds, \quad (2.100)$$

where again τ denotes the tangential vector of the curve C . Noting that $\nabla_z \left(\frac{1}{\|y-z\|} \right) = \frac{y-z}{\|y-z\|^3}$, we can now compute

$$\begin{aligned}
\varphi_C^{\text{TMS}}(v) &= -f'(t) \cdot \frac{\mu_0}{4\pi} \cdot \int_{\Omega} \langle \sigma \int_F \eta \times \frac{y-z}{\|y-z\|^3} dS(z), \nabla v(y) \rangle dV(y) \\
&= -f'(t) \cdot \frac{\mu_0}{4\pi} \cdot \int_{\Omega} \int_F \langle \sigma \cdot \left(\eta \times \frac{y-z}{\|y-z\|^3} \right), \nabla v(y) \rangle dS(z) dV(y) \\
&\stackrel{\text{Fubini}}{=} -f'(t) \cdot \frac{\mu_0}{4\pi} \cdot \int_F \int_{\Omega} \langle \sigma \cdot \left(\eta \times \frac{y-z}{\|y-z\|^3} \right), \nabla v(y) \rangle dV(y) dS(z) \\
&= -f'(t) \cdot \frac{\mu_0}{4\pi} \cdot \int_F \int_{\Omega} \langle \sigma \nabla v(y) \times \frac{z-y}{\|z-y\|^3}, \eta \rangle dV(y) dS(z) \\
&= f'(t) \cdot \psi_C^{\text{MEG}}(v).
\end{aligned}$$

From equation (2.99), we thus finally get

$$-f'(t) \cdot \Phi_B^{\text{S}}(C) = -\varphi^{\text{dip}}(u_{\varphi_C^{\text{TMS}}}) = \langle M, -\nabla u_{\varphi_C^{\text{TMS}}}(x_0) \rangle = \langle E^{\text{S}}(x_0, t), M \rangle.$$

This shows the MEG/TMS flux reciprocity law for the secondary magnetic flux (2.98).

For the primary flux, we can similarly compute

$$\begin{aligned}
-f'(t) \cdot \Phi_B^{\text{P}}(C) &= -f'(t) \cdot \frac{\mu_0}{4\pi} \cdot \int_F \langle M \times \frac{z-x_0}{\|z-x_0\|^3}, \eta \rangle dS(z) \\
&= -f'(t) \cdot \frac{\mu_0}{4\pi} \cdot \langle \int_F \eta \times \frac{x_0-z}{\|x_0-z\|^3} dS(z), M \rangle \\
&\stackrel{\text{Eq. (2.100)}}{=} -f'(t) \cdot \frac{\mu_0}{4\pi} \cdot \langle \int_C \frac{\tau}{\|x_0-z\|} ds(z), M \rangle \\
&\stackrel{\text{Eq. (2.95)}}{=} \langle -f'(t) \cdot A^{\text{S}}(x_0), M \rangle = \langle -\frac{\partial A}{\partial t}(x_0, t), M \rangle \\
&= \langle E^{\text{P}}(x_0, t), M \rangle.
\end{aligned}$$

This shows the MEG/TMS flux reciprocity law for the primary magnetic flux (2.97). But this also immediately implies the MEG/TMS reciprocity law for the total magnetic flux (2.96), since

$$\begin{aligned}
-f'(t) \cdot \Phi_B(C) &= -f'(t) \cdot \Phi_B^{\text{P}}(C) + (-f'(t)) \cdot \Phi_B^{\text{S}}(C) = \langle E^{\text{P}}(x_0, t), M \rangle + \langle E^{\text{S}}(x_0, t), M \rangle \\
&= \langle E(x_0, t), M \rangle.
\end{aligned}$$

□

2.9.4 Relation of reciprocity to transfer matrix approaches

Above, we already noted the similarity of reciprocity-based forward simulations to transfer matrix approaches. In this section, we want to elaborate on this. More concretely, we will show that transfer matrix approaches are exactly the same as expressing the reciprocity theorem 2.9.2 in a coordinate system. We begin by introducing the common formulation of transfer matrix approaches and refer to [238] for details.

Transfer matrices

A typical finite element approach consists of specifying a *finite-dimensional* vector space H , a bilinear form $a : H \times H \rightarrow \mathbb{R}$, and a linear functional $\varphi : H \rightarrow \mathbb{R}$, and searching for a solution u to the problem

$$\text{Find } u \in H \text{ such that for all } v \in H \text{ we have } a(u, v) = \varphi(v). \quad (2.101)$$

As above, we will assume a to be symmetric. Since H is, by assumption, finite-dimensional, we can find a basis $\mathfrak{B} = \{u_1, \dots, u_N\}$ of H . Then, any element $v \in H$ can be written in a unique way as $v = \sum_{j=1}^N y_j \cdot u_j$ for $y_1, \dots, y_N \in \mathbb{R}$. This implies that $u \in H$ solves problem (2.101) if, and only if, $a(u, u_i) = \varphi(u_i)$ for all $1 \leq i \leq N$. Thus, $u = \sum_{j=1}^N x_j \cdot u_j \in H$ solves (2.101) if, and only if,

$$\sum_{j=1}^N a(u_j, u_i) \cdot x_j = \varphi(u_i) \text{ for all } 1 \leq i \leq N$$

If we now define a matrix $A = (a_{ij})_{1 \leq i, j \leq N} \in \mathbb{R}^{N \times N}$ via $a_{ij} := a(u_j, u_i)$ and a vector $b = (b_i)_{1 \leq i \leq N} \in \mathbb{R}^N$ via $b_i := \varphi(u_i)$, we see that $u = \sum_{j=1}^N x_j \cdot u_j \in H$ solves (2.101) if, and only if, the vector $x = (x_j)_{1 \leq j \leq N}$ solves the linear equation $A \cdot x = b$. Additionally, note that A is symmetric since a is symmetric.

As already described above, we are typically not interested in the complete FEM solution u , but only in some measurement $\psi(u)$ of the solution u , which can be described by a linear functional $\psi : H \rightarrow \mathbb{R}$. Now define a vector $w = (w_i)_{1 \leq i \leq N} \in \mathbb{R}^N$ via $w_i := \psi(u_i)$. Now, let $u = \sum_{j=1}^N x_j \cdot u_j \in H$ be a solution of (2.101). Then

$$\psi(u) = \sum_{j=1}^N \psi(u_j) \cdot x_j = w^\top \cdot x.$$

If now $t_\psi \in \mathbb{R}^N$ is a solution of $A \cdot t_\psi = w$, we have

$$\psi(u) = w^\top \cdot x = (A \cdot t_\psi)^\top \cdot x = t_\psi^\top \cdot A \cdot x = t_\psi^\top \cdot b. \quad (2.102)$$

We call t_ψ^\top the *transfer matrix row corresponding to the functional ψ* . Note that, once the transfer matrix row t_ψ^\top is known, we can simulate the measurement $\psi(u)$ due to an *arbitrary source term* by simply assembling its FEM right-hand side and computing a vector-vector product. In particular, we do not need to solve an additional linear system for each source. Note, however, that the transfer matrix row t_ψ^\top is itself computed by solving the linear system $A \cdot t_\psi = w$. But in applications where we have much fewer measurement sensors than sources, this can massively decrease computation times, as we do not need to solve a linear system for each source, but only for each sensor.⁴¹

We want to emphasize that we did not assume A to be invertible. In many applications, such as the EEG, the MEG, the tES, or the TMS forward problems, the matrix A is in fact not invertible. A consequence of this is that the transfer matrix row t_ψ^\top is not uniquely determined, and can be any of the (potentially infinitely many) solutions of $A \cdot t_\psi = w$.

⁴¹One example of such an application is the computation of EEG and MEG lead fields, where one typically has some hundred sensors and tens of thousands of sources.

Relation to reciprocity

Above, we claimed that this transfer matrix approach is nothing other than expressing the general reciprocity theorem 2.9.2 in a coordinate system. To make this precise, we introduce some notation.

As noted above, any element $y \in H$ can be uniquely expressed in the basis \mathfrak{B} as $y = \sum_{i=1}^N y_i \cdot u_i$. We call the vector $[y]_{\mathfrak{B}} := (y_i)_{1 \leq i \leq N} \in \mathbb{R}^N$ the *coordinates of y in the basis \mathfrak{B}* . Furthermore, if $\varphi : H \rightarrow \mathbb{R}$ is a linear functional, it is already uniquely determined from its values on the basis \mathfrak{B} . We write $[\varphi]_{\mathfrak{B}} := (\varphi(u_i))_{1 \leq i \leq N} \in \mathbb{R}^N$ for the vector containing these values. Now, note that

$$\varphi(y) = \sum_{i=1}^N \varphi(u_i) \cdot y_i = \langle [\varphi]_{\mathfrak{B}}, [y]_{\mathfrak{B}} \rangle. \quad (2.103)$$

In 2.9.4, we saw that the finite element problem

$$\text{Find } u \in H \text{ such that for all } v \in H \text{ we have } a(u, v) = \varphi(v).$$

is equivalent to the linear equation $A \cdot x = b$, where, using the notation from this subsection, we have $x = [u]_{\mathfrak{B}}$ and $b = [\varphi]_{\mathfrak{B}}$. Written more succinctly, this means that u_{φ} is a FEM solution due to the source functional φ if, and only if, it fulfills the linear system $A \cdot [u_{\varphi}]_{\mathfrak{B}} = [\varphi]_{\mathfrak{B}}$.

Now, let $\psi : H \rightarrow \mathbb{R}$ denote a measurement functional, which as a source functional generates the FEM solution u_{ψ} . Then, for an arbitrary source functional $\varphi : H \rightarrow \mathbb{R}$ with corresponding FEM solution u_{φ} , we have by the reciprocity theorem 2.9.2 and by equation (2.103) that

$$\psi(u_{\varphi}) = \varphi(u_{\psi}) = \langle [u_{\psi}]_{\mathfrak{B}}, [\varphi]_{\mathfrak{B}} \rangle = [u_{\psi}]_{\mathfrak{B}}^T \cdot [\varphi]_{\mathfrak{B}}.$$

By definition, u_{ψ} is given as a FEM solution corresponding to the functional ψ acting as a source, which by the derivation above is equivalent to the linear system $A \cdot [u_{\psi}]_{\mathfrak{B}} = [\psi]_{\mathfrak{B}}$. But since $A \cdot t = [\psi]_{\mathfrak{B}}$ is also the defining equation of a transfer matrix row for the functional ψ , we see that *a transfer matrix row for the functional ψ is exactly the same as the coordinate representation of a FEM solution corresponding to the functional ψ acting as a source*. Written more compactly, we have

$$t_{\psi} = [u_{\psi}]_{\mathfrak{B}}.$$

Furthermore, we see that multiplication of a transfer matrix row t_{ψ}^T with the FEM right-hand side $[\varphi]_{\mathfrak{B}}$ is simply the coordinate representation of evaluating the functional φ in the FEM solution u_{ψ} , which by the reciprocity theorem 2.9.2 computes the evaluation of the functional ψ in the FEM solution u_{φ} , i.e. the quantity we are really interested in. This is what we mean by the statement that transfer matrix approaches are just reciprocity expressed in coordinates.

This finishes our discussion on the reciprocity laws of quasi-static bioelectromagnetism. Moreover, this also finishes our discussion of the forward problems of quasi-static bioelectromagnetism. We will now jump to the inverse problem of EEG and MEG.

Chapter 3

The EEG and MEG inverse problems

In the preceding chapter, we discussed how to simulate the electric and magnetic fields due to a given neural activity. In this chapter, we want to discuss the opposite direction, i.e., the estimation of neural activity from non-invasively measured EEG and MEG signals. The process of using EEG and MEG measurements to infer information about the underlying neural activity is called the inverse problem of EEG and MEG. Since there are non-zero current densities where *both* the generated electric field and the generated magnetic field vanish outside of the head domain, this problem is fundamentally ill-posed [83]. Any approach for solving the inverse problem thus has to make some prior assumptions about the source activity it wants to reconstruct.

Typically, the first assumption is that the neural activity can be adequately modeled by a linear combination of so-called point dipoles [40], see also section 2.5. A point dipole is given by a position in the head, say $x_0 \in \mathbb{R}^3$, and so-called dipole moment, say $M \in \mathbb{R}^3$, encoding the orientation and intensity of the neural current. Mathematically, the current density of the corresponding point dipole is then described by the distribution $M \cdot \delta_{x_0}$. The goal of the EEG and MEG inverse problem is now to explain the measured data using such dipoles.

To this end, a plethora of different strategies, based on different assumptions and with different interpretations of the generated output, have been suggested [127, 77].

One common such assumption is that the measured data can be well explained by a small number of dipoles, or even just one dipole. This is, e.g., frequently used in the presurgical evaluation of pharmacoresistant focal epilepsy patients, where dipole reconstructions aid in the planning of brain resections [164]. Another well-known application where this assumption is commonly used is the study of evoked responses, where subjects are presented with some form of stimuli, e.g., auditory or somatosensory, and the resulting brain response is recorded [6, 132]. An inverse approach tailored to this kind of scenario is the dipole scan [68]. In this approach, one iterates over a typically large set of possible dipole positions. For each such position, one then checks how well a dipole at this position can explain the measured data by performing a simple least squares fit to the data. Positions with a good fit are then designated as plausible generators of the measured data.

When no prior assumptions about the number of sources can be made, one typically uses a

so-called distributed inverse method [31]. Here, one specifies a set of possible dipolar source locations in the brain, and then tries to simultaneously estimate the dipole moments at each source position. Since this problem is typically heavily underdetermined, one has to specify some heuristic for choosing one of the many different source configurations that explain the data equally well. The classical choice is the minimum norm solution of the corresponding regularized least squares problem, which in the biomagnetism community is called minimum norm estimation (MNE) [94]. It is, however, well-known that MNE is biased towards superficial sources [119], which is simply due to the fact that superficial sources are closer to the measurement sensors than sources deep inside the brain, and can hence explain the measured data with smaller dipole moments than deep sources. A lot of approaches have been suggested to overcome this drawback of the MNE method by introducing a depth-dependent weighting into the reconstruction, e.g. [67, 38]. Another massively successful such approach, which is in particular focus in this chapter, is the standardized low resolution brain electromagnetic tomography (sLORETA) [149]. The sLORETA method is derived by assuming a simple noise model consisting of neural background activity and sensor noise. In this model, the author then computes variances of inverse source estimations of pure noise signals, which are then interwoven with the minimum norm estimations of the neural activity to assign a score to each source position. Neural activity is then estimated by looking for large values of this score. One of the appealing properties of the sLORETA approach is that it is able to correctly localize neural activity originating from a single dipolar source in a noiseless scenario, which is not true for the MNE approach and its other derivatives. While in the original paper [149] this property was only shown empirically, later works also gave rigorous mathematical proofs of this fact [192, 78]. Building on top of these results, the author of the original paper later published [150], where he generalizes the sLORETA approach to a whole family of inverse approaches, each member of which possesses the ability to correctly localize single dipolar sources in noiseless scenarios. In particular, [150] introduced the so-called eLORETA approach, which has also seen widespread use in the bioelectromagnetism community.

In this chapter, we will show that the family of reconstruction approaches defined in [150] is exactly the same as the family of dipole scanning approaches. This means that, while in the literature, sLORETA and eLORETA are always introduced and interpreted as distributed reconstruction approaches, the scores they use to localize neural activity are, at each dipole position, mathematically the same as the fit values of a single dipole at this position to the data. Once this result is derived, it is almost trivial to see why sLORETA is able to correctly localize noiseless dipolar signals. Furthermore, we will show how, using this identification, various results regarding the biasedness of the sLORETA and eLORETA approaches distributed over the literature can be unified as special cases of one natural result regarding the biasedness of dipole scans. Another particularly interesting consequence of this identification is that it enables us to directly estimate source orientations using sLORETA, where previously it was thought that sLORETA could only estimate source positions, but not orientations [223, 77].

Yet another framework to approach the EEG and MEG inverse problems is through so-called *beamformers*. While these methods originate from radar and sonar applications, they were later introduced to bioelectromagnetism in [210, 209]. The general idea is to create a so-called *adaptive spatial filter*. Given some target region, this filter is supposed to be “tuned” towards the target region in such a way that it lets signals originating from the target region pass through, while filtering out signals originating from non-target regions. Sources in the brain are then estimated by dividing the brain into different target regions. For each such region, one computes a corresponding adaptive spatial filter and applies it to the measured signals. Source location estimation then (typically) proceeds by computing the power of the filtered signal and looking for brain re-

gions where this power is large. If the filter can be assumed to properly attenuate contributions from non-target regions, it follows that the power in the post-filter signal measures the neural activity in the targeted area, and can thus be used to localize active brain regions [190]. Different realizations of this idea have been suggested in the literature. The most well-established of these are so-called unit-gain (UG) beamformers, array-gain (AG) beamformers, unit-noise-gain (UNG) beamformers [190], neural activity index (NAI) beamformers [210], and synthetic aperture magnetometry (SAM) beamformers [222]. An overview of the current state-of-the-art of beamformer methodology in bioelectromagnetism research can be found in [231].

In this chapter, we will show that in scenarios where the measurement data can be described by a single dipolar source with additive uncorrelated noise, the beamformers mentioned above, with the notable exception of the UG beamformer, are equivalent to dipole scanning approaches. This identification has a number of interesting consequences. First of all, similar to the sLORETA case, it enables an almost trivial proof that the beamformers mentioned above, again except the UG beamformer, are able to correctly localize single dipolar sources. While this result is already known in the literature, its proofs previously consisted of arduous computations, where it was difficult to understand why this result is true [192, 78, 135]. Furthermore, since both AG and UNG beamformers are equivalent to dipole scanning in single-source scenarios, they are, in particular, also equivalent to one another. This gives a theoretical explanation for previously observed empirical results, such as those in our own publication [30]. Furthermore, we will use the methods derived in the identification of dipole scanning and beamforming to show that the suggested vector version of the NAI in [210], as well as the vector beamformers introduced in [190], will, in general, lead to a biased source estimation.

Additionally, we have already mentioned above that the UG beamformer is not equivalent to a dipole scan. In fact, it is well known that the UG beamformer in general fails to localize single dipolar sources [190]. In the literature, this is typically attributed to a *depth bias*, or *center of head bias* [137, 231]. If one is instead only interested in reconstructing the orientation of a neural source with an already known location, it was not yet clear how well the UG beamformer performed. This was investigated by us in [30], where we showed that, even in the case that one is only interested in the orientation, the UG beamformer introduces a bias into the reconstruction. Nevertheless, we were able to show that, as long as the SNR is high enough, the UG beamformer will still yield an accurate result. In this chapter, we will present these investigations.

Another result from [30] that we want to discuss lies in the combination of EEG and MEG. As the MEG is insensitive to “radial” source orientations, i.e., those sources whose moment is oriented roughly normal to the skull surface, one typically cannot expect to reliably reconstruct the radial component of a source orientation from a MEG measurement [92, 153]. The EEG, on the other hand, is sensitive to radial orientation components, but is generally less sensitive to tangential orientation components than the MEG [153, 74]. This suggests combining the complementary strengths of the EEG and the MEG by, e.g., first computing an orientation estimate from the EEG alone, and then refining the tangential component of this estimate by utilizing the MEG. We will show that such an approach is unbiased, i.e., that removing the radial component from the beamforming reconstruction does not introduce a bias into the orientation estimation for the NAI and SAM beamformers.

The chapter is organized as follows. In section 3.1, we will first introduce the necessary notations and definitions. In section 3.2, we will then derive the equivalence between dipole scanning and sLORETA, and discuss how this identification aids in the interpretation of the sLORETA method. In section 3.3, we will then investigate the relation between beamforming and dipole scanning.

3.1 Notations and Definitions

3.1.1 Lead fields

If one wants to solve the EEG and MEG inverse problems, one first needs to solve the corresponding forward problems. We refer to chapter 2 of this thesis for an in-depth discussion on how to solve the forward problems. We now want to fix the notation for the forward solutions that is used during this chapter.

Definition 3.1.1. Let $x_0 \in \mathbb{R}^3$ denote a dipole position and $M \in \mathbb{R}^3$ a dipole moment. Let L denote the number of sensors. We then denote by $l_{x_0, M} \in \mathbb{R}^L$ the vector containing the sensor measurements a dipole at position x_0 with moment M would generate, and call this vector the lead field of the dipole $M \cdot \delta_{x_0}$. If we denote by $e_i \in \mathbb{R}^3$ the i -th unit vector, we then define the lead field L_{x_0} at position x_0 to be the matrix

$$L_{x_0} := (l_{x_0, e_1}, l_{x_0, e_2}, l_{x_0, e_3}) \in \mathbb{R}^{L \times 3}.$$

If we have discretized the set of possible dipole positions in the brain to a set of k positions x_1, \dots, x_k , we define the complete lead field L to be

$$L := (L_{x_1}, L_{x_2}, \dots, L_{x_k}) \in \mathbb{R}^{L \times (3 \cdot k)}.$$

We call such a set of positions a source space, and refer to section 2.7.3 for a discussion on how to construct such a set.

If we want to explicitly talk about EEG measurements, we indicate this by adding a superscript, i.e., $l_{x_0, M}^E, L_{x_0}^E, L^E$, and call these EEG lead fields. Similarly for MEG measurements, we write $l_{x_0, M}^M, L_{x_0}^M, L^M$ and speak of MEG lead fields.

The EEG lead field is column-wise only determined up to an additive constant, as, e.g., discussed in chapter 2. To remove this ambiguity, one typically chooses from the set of valid EEG lead fields the unique representative with the property that the average of the lead field vector entries is zero, and calls this common average referencing. In the following, we will always assume that the EEG lead field is common average referenced. If we define $\mathbb{R}_{\text{CAR}}^L := \{x \in \mathbb{R}^L \mid \sum_{i=1}^L x_i = 0\}$, this means that we always assume $l_{x_0, M}^E \in \mathbb{R}_{\text{CAR}}^L$ for any dipole position x_0 and moment M .

In the following, we will always treat the EEG and MEG cases separately, i.e., a lead field will always be either an EEG or an MEG lead field. Because EEG lead fields are always contained in the space $\mathbb{R}_{\text{CAR}}^L$, we will later on need slightly different definitions in the EEG case than in the MEG case.

Definition 3.1.2. In the EEG case, we call $\mathbb{R}_{\text{CAR}}^L$ the lead field range. Furthermore, we call a matrix $C \in \mathbb{R}^{L \times L}$ positive definite on the lead field range if C is symmetric, $C \cdot \mathbb{R}_{\text{CAR}}^L = \mathbb{R}_{\text{CAR}}^L$, and furthermore $C|_{\mathbb{R}_{\text{CAR}}^L} : \mathbb{R}_{\text{CAR}}^L \rightarrow \mathbb{R}_{\text{CAR}}^L$ is a positive definite operator with respect to the standard inner product on $\mathbb{R}_{\text{CAR}}^L$. In the MEG case, we call \mathbb{R}^L the lead field range, and if we call a matrix $C \in \mathbb{R}^{L \times L}$ positive definite on the lead field range, we simply mean it to be a symmetric positive definite matrix.

Note that in both the EEG and the MEG case, this definition implies that if x_0 is a dipole position with $\text{rank}(L_{x_0}) = 3$, and $C \in \mathbb{R}^{L \times L}$ is positive definite on the lead field range, the matrix $L_{x_0}^T C L_{x_0}$ is positive definite, and in particular invertible.

Finally, note that since the electric potential and the magnetic field depend linearly on the neural activity, we have $l_{x_0, M} = L_{x_0} \cdot M$.

3.1.2 Dipole scanning

We now want to introduce dipole scanning approaches. As mentioned in the introduction, the underlying assumption of a dipole scanning approach is that the data can be explained using a single dipole. To compare measurements and simulations, one first needs to define a metric. In dipole scanning approaches, one chooses a metric derived from an inner product. If one denotes the standard inner product on \mathbb{R}^L by $\langle \cdot, \cdot \rangle$, the set of all inner products on the lead field range is parametrized by positive definite operators C with respect to this standard inner product on the lead field range, where each such C defines a norm via

$$\|v\|_C := \sqrt{\langle C \cdot v, v \rangle}.$$

Assume now that we have a measurement vector $d \in \mathbb{R}^L$. In the EEG case, we assume that d is common average referenced. For a candidate dipole position x_0 , one then computes a minimizer

$$j_C(d, L_{x_0}) := \arg \min_{j \in \mathbb{R}^3} \|d - L_{x_0} \cdot j\|_C^2, \quad (3.1)$$

which can then be used to compute the *residual variance*

$$\text{rv}_C(d, L_{x_0}) := \|d - L_{x_0} \cdot j_C(d, L_{x_0})\|_C^2. \quad (3.2)$$

One then typically reports the *goodness of fit*

$$\text{GOF}_C(d, L_{x_0}) := 1 - \frac{\text{rv}_C(d, L_{x_0})}{\|d\|_C^2}, \quad (3.3)$$

and interprets positions with a higher goodness of fit (GOF) as more likely to be the generator of the measured data. Note that

$$0 \leq \min_{j \in \mathbb{R}^3} \|d - L_{x_0} \cdot j\|_C^2 \leq \|d - L_{x_0} \cdot 0\|_C^2 = \|d\|_C^2.$$

This implies that $0 \leq \text{rv}_C(d, L_{x_0}) \leq \|d\|_C^2$, and $0 \leq \text{GOF}_C(d, L_{x_0}) \leq 1$. By, on the one hand, looking at the case where d is orthogonal to the lead field columns, and, on the other hand, the case where d is contained in the lead field range, we see that these inequalities are sharp.

To perform a dipole scan, one then constructs a source space x_1, \dots, x_k of possible dipole positions. For each position in the source space, one then computes the corresponding goodness of fit. The location of the dipole is then estimated by looking at positions where the goodness of fit is large.

If L_{x_0} does not have full rank, there is no unique minimizer of equation (3.1). Note, however, that in this case both the residual variance as well as the goodness of fit are still well-defined, as they are independent of the choice of minimizer of equation (3.1).

3.1.3 (Generalized) sLORETA

The sLORETA method was originally published in [149]. The author then later published [150], where he introduced a whole family of reconstruction approaches, of which the classical sLORETA method from [149] is one particular member. We will now introduce this family of reconstruction approaches.

Each member of the family of generalized sLORETA approaches is parametrized by a matrix $C \in \mathbb{R}^{L \times L}$. In [150], the author only requires this matrix to be symmetric. In the following, we

will additionally assume that C is positive definite on the lead field range. We argue that the author in [150] should also have required C to be positive definite on the lead field range, since the formulation of the method requires that, for each lead field L_{x_0} , the matrix $(L_{x_0}^\top C L_{x_0})^{-\frac{1}{2}}$ is well defined, and there is no reason for this to be the case without the positive definiteness condition. Furthermore, the concrete choices for C presented in [150] are always positive definite on the lead field range.

Definition 3.1.3 (Generalized sLORETA). Let $C \in \mathbb{R}^{L \times L}$ be positive definite on the lead field range. Let $L_{x_0} \in \mathbb{R}^{L \times 3}$ be the lead field at position x_0 . Then, the generalized sLORETA estimation of the neural activity is given by

$$j_C^{\text{sLORETA}}(d, L_{x_0}) = (L_{x_0}^\top C L_{x_0})^{-\frac{1}{2}} L_{x_0}^\top C d \in \mathbb{R}^3. \quad (3.4)$$

To estimate the location of neural activity, one then constructs a source space x_1, \dots, x_k and computes for each position x_i the vector $j_C^{\text{sLORETA}}(d, L_{x_i})$. The locations of neural sources are then estimated by looking for large values of $\|j_C^{\text{sLORETA}}(d, L_{x_i})\|_2^2$, where $\|\cdot\|_2$ denotes the euclidean norm on \mathbb{R}^3 .

Now, let x_1, \dots, x_k denote a source space, with corresponding complete lead fields L^E and L^M as defined in definition 3.1.1. Furthermore, let p_{CAR} denote the orthogonal projection of \mathbb{R}^L onto $\mathbb{R}_{\text{CAR}}^L$ with respect to the standard inner product. The original sLORETA method, as defined in [149], then corresponds to the choice

$$C = \left(L^E (L^E)^\top + \alpha \cdot p_{\text{CAR}} \right)^+ \quad (3.5)$$

in the EEG case, and

$$C = \left(L^M (L^M)^\top + \alpha \cdot \text{Id}_L \right)^{-1} \quad (3.6)$$

in the MEG case, where “+” denotes the pseudo-inverse and $\alpha \geq 0$ is a user-specified parameter. The eLORETA approach, as it was introduced in [150], corresponds to the choice

$$C = \left((L^E) (W^E)^{-1} (L^E)^\top + \alpha \cdot p_{\text{CAR}} \right)^+ \quad (3.7)$$

in the EEG case, and

$$C = \left((L^M) (W^M)^{-1} (L^M)^\top + \alpha \cdot \text{Id}_L \right)^{-1} \quad (3.8)$$

in the MEG case, where $W^E = \text{diag}(W_1^E, \dots, W_k^E) \in \mathbb{R}^{(3k) \times (3k)}$ and $W^M = (W_1^M, \dots, W_k^M) \in \mathbb{R}^{(3k) \times (3k)}$, $W_i^E, W_i^M \in \mathbb{R}^{3 \times 3}$, are weight matrices defined by the equations

$$W_i^E = \left((L_{x_i}^E)^\top \left((L^E) (W^E)^{-1} (L^E)^\top + \alpha \cdot p_{\text{CAR}} \right)^+ (L_{x_i}^E) \right)^{\frac{1}{2}} \quad \text{for } 1 \leq i \leq k \quad (3.9)$$

$$W_i^M = \left((L_{x_i}^M)^\top \left((L^M) (W^M)^{-1} (L^M)^\top + \alpha \cdot \text{Id}_L \right)^{-1} (L_{x_i}^M) \right)^{\frac{1}{2}} \quad \text{for } 1 \leq i \leq k \quad (3.10)$$

The author in [150] then suggests finding such weight matrices W^E and W^M by simply performing a fixed point iteration on equations (3.9) and (3.10), with the identity matrix as an initial

guess. The author does, to the best of our knowledge, not prove the convergence of this fixed point iteration, nor whether there even exists a solution to the equations (3.9) and (3.10). The convergence of the iteration is instead simply postulated.

eLORETA can be motivated as follows. We will only discuss the MEG case.¹ We will suppress the “M” superscript for this computation. Imagine α as a regularization parameter. In the limit $\alpha \rightarrow 0$, we get for C that

$$C = (L \cdot W^{-1} \cdot L^\top)^{-1}. \quad (3.11)$$

But then, equation (3.10) yields

$$W_i = \left(L_{x_i}^\top (LW^{-1}L)^{-1} L_{x_i} \right)^{\frac{1}{2}} = \left(L_{x_i}^\top C L_{x_i} \right)^{\frac{1}{2}} \iff W_i^{-1} = \left(L_{x_i}^\top C L_{x_i} \right)^{-\frac{1}{2}}. \quad (3.12)$$

If we now gather all eLORETA vectors in one large vector

$$j^{\text{eLORETA}} = \begin{pmatrix} j_C^{\text{sLORETA}}(d, L_{x_1}) \\ \vdots \\ j_C^{\text{sLORETA}}(d, L_{x_k}) \end{pmatrix} \in \mathbb{R}^{3 \cdot k},$$

we get

$$\begin{aligned} L \cdot j^{\text{eLORETA}} &\stackrel{(3.4)}{=} L \cdot \text{diag}((L_{x_1}^\top C L_{x_1})^{-\frac{1}{2}}, \dots, (L_{x_k}^\top C L_{x_k})^{-\frac{1}{2}}) \cdot L^\top \cdot C \cdot d \\ &\stackrel{(3.12)}{=} L \cdot W^{-1} \cdot L^\top \cdot C \cdot d \\ &\stackrel{(3.11)}{=} d. \end{aligned}$$

Hence, the combination of eLORETA vectors j^{eLORETA} can be interpreted as a selection of dipole moments that together *exactly* reproduce the measurement data d . In the language of the bioelectromagnetism community, this means that the eLORETA vectors, at least in the limit of $\alpha \rightarrow 0$, form an exact current density reconstruction of the measurement data. This is the origin of the name *exact LORETA*, or *eLORETA* for short.

Finally, we want to note that in both [149] and [150] the author differentiates between the cases of “known” and “unknown” orientation. What is meant by this is that, in bioelectromagnetism, one oftentimes assumes, based on physiological reasoning, that the neural currents are flowing normal to the cortical surface [83]. If one now has a high confidence in one’s segmentation of the brain and its surfaces, one can at each source position x estimate its normal direction, say η , and only allow sources at position x to be oriented along the direction η . One can then replace each occurrence of $L_x \in \mathbb{R}^{L \times 3}$ above with $l_{x,\eta} \in \mathbb{R}^{L \times 1}$. This distinction makes no difference for the identification of the generalized sLORETA approaches with the class of dipole scanning approaches that we will present below. In fact, in the “scalar” case, i.e., when using $l_{x,\eta}$ instead of L_x , it is even possible to give a simpler argument than the one we are going to present. Our argument below, however, immediately carries over to the scalar case, and we will hence not give a separate discussion of the scalar case in this thesis.

3.1.4 Beamforming

We follow [135] with slight modifications. Let L again denote the number of sensors and let Ω denote the head domain. Assume that the signals originating from some subset $\Omega_0 \subset \Omega$ of the

¹For the EEG case, the exact same logic applies. We only restrict ourselves to the MEG case to keep the presentation more succinct.

brain can be described as being generated by a linear combination of dipolar sources at positions x_1, \dots, x_n with moments M_1, \dots, M_n .² Denote by

$$L_n := (l_{x_1, M_1}, \dots, l_{x_n, M_n}) \in \mathbb{R}^{L \times n}$$

their combined lead field. In the following, we assume $n < L$ and $\text{rank}(L_n) = n$. As described in the introduction, the goal is to create a filter that lets the contributions originating from the chosen sources through, but filters out the remaining activity. In beamforming approaches, such a filter is given by a matrix $W \in \mathbb{R}^{L \times n}$, where the filter output of a signal $d \in \mathbb{R}^L$ is given by $W^\top \cdot d \in \mathbb{R}^n$. Since we do not know what other sources might be active, we cannot explicitly enforce the filter properties of W . Instead, the approach is to make W adaptively learn from the data what other sources are present, and how they can be suppressed. Because of this, beamformers are also sometimes called adaptive spatial filters.

This is done as follows. Let d denote the measurement vector, which we now assume to be an \mathbb{R}^L -valued random vector contaminated by noise. Then, the filter $W_M(L_n)$ is defined by

$$W_M(L_n) := \arg \min_{W^\top \cdot L_n = M^\top} \mathbb{E} [\|W^\top \cdot d\|_2^2], \quad (3.13)$$

where $M \in \mathbb{R}^{n \times n}$ is some pre-specified matrix. Here, the side constraint $W^\top \cdot L_n = M^\top$ ensures that signals originating from the dipoles of interest are not filtered out, while minimizing the expected power of the post-filter signal is expected to filter out contributions not originating from the sources of interest. Furthermore, different beamformers correspond to different choices of M . In the following, we always assume $\text{rank}(M) = n$. If we now define

$$R := \mathbb{E} [d \cdot d^\top] \in \mathbb{R}^L$$

to be the covariance matrix³, we can compute

$$\mathbb{E} [\|W^\top \cdot d\|_2^2] = \mathbb{E} [(W^\top \cdot d)^\top (W^\top \cdot d)] = \mathbb{E} [d^\top W W^\top d] = \text{Tr} (W^\top R W), \quad (3.14)$$

where the last equation follows by just writing out the matrix products. In the following, we will always assume that R is positive definite on the lead field range and has the orthogonal complement of the lead field range as its kernel, and will denote by R^{-1} an arbitrary matrix that is positive definite on the lead field range, and acts as an inverse of R on the lead field range. Such a matrix is, e.g., given by the pseudoinverse of R .⁴

²In the context of somatosensory stimulation, one might, e.g., take Ω_0 to be a subset of the somatosensory cortex, and model the signal as being generated by a single dipole. For auditory evoked potentials, one might, e.g., take Ω_0 to be the union of two regions in the left and right auditory cortex, and assume the signal is generated by two dipoles. Yet another example would be to take Ω_0 to be a small subset of the brain, which might contain neural sources of different orientations. One might then decide to employ a so-called *regional dipole*, where one models any activity from this region as being generated from a dipole with fixed “equivalent” position, say x , but with a freely varying orientation [181]. In the notation above, this means that one chooses $n = 3$, $x_1 = x_2 = x_3 = x$, and $M_1 = (1, 0, 0)$, $M_2 = (0, 1, 0)$, and $M_3 = (0, 0, 1)$.

³Note that this is really the second moment matrix, but it has become a convention in the bioelectromagnetism community to call this matrix the covariance matrix nonetheless, see e.g. [190]. In this thesis, we will follow this convention.

⁴We decided to follow the literature and use the symbol R^{-1} for this matrix, even though R will in general, e.g., in the EEG case, *not* be an invertible matrix. We think this is justified, as R only enters as a linear map acting on the lead field range, and we do not care what R does on the orthogonal complement of the lead field range. From a mathematical perspective, it would probably be more natural to develop the theory in some abstract Hilbert space H and then specialize for the EEG to the case $H = \mathbb{R}_{\text{CAR}}^L$. We decided against this, as this would make it significantly more difficult to compare our results to the wider beamforming literature, especially if the reader does not have a lot of experience with mathematics.

Lemma 3.1.4. *The optimization problem (3.13) has a unique solution whose columns are contained in the lead field range. This solution is given by*

$$W_M(L_n) = R^{-1}L_n (L_n^\top R^{-1}L_n)^{-1} M. \quad (3.15)$$

Proof. Let $W = (w_1, \dots, w_n) \in \mathbb{R}^{L \times n}$, $L = (l_1, \dots, l_n) \in \mathbb{R}^{L \times n}$, and $M = (m_1, \dots, m_n) \in \mathbb{R}^{n \times n}$. Then

$$W^\top RW = \begin{pmatrix} \langle w_1, Rw_1 \rangle & \dots & \langle w_1, Rw_n \rangle \\ \vdots & \ddots & \vdots \\ \langle w_n, Rw_1 \rangle & \dots & \langle w_n, Rw_n \rangle \end{pmatrix},$$

and thus

$$\text{Tr}(W^\top RW) = \sum_{i=1}^n \langle w_i, Rw_i \rangle.$$

Furthermore, we have for the side constraint that

$$W^\top \cdot L = M^\top \iff L^\top w_1 = m_1, \dots, L^\top w_n = m_n.$$

We thus see that the optimization problem 3.13 for

$$W_M(L_n) = (w_M^1(L_n), \dots, w_M^n(L_n)) \in \mathbb{R}^{L \times n}$$

is equivalent to the set of n independent optimization problems

$$w_M^i(L_n) = \arg \min_{L^\top w = m_i} \langle w, Rw \rangle \quad \text{for } 1 \leq i \leq n. \quad (3.16)$$

Let $1 \leq i \leq n$ be fixed. Since the lead field range is an invariant subspace with respect to R , and the orthogonal complement of the lead field range is the kernel of R , we have

$$\min_{L^\top w = m_i} \langle w, Rw \rangle = \min_{\substack{L^\top w = m_i \\ w \in \text{Im}(R)}} \langle w, Rw \rangle$$

It thus suffices to search for minimizers of (3.16) in the space

$$\{w \in \mathbb{R}^L | L^\top w = m_i\} \cap \text{Im}(R) = \{w \in \mathbb{R}^L | L^\top w = m_i\} \cap \text{Ker}(R)^\perp.$$

Since R is by assumption positive definite on $\text{Im}(R)$, it follows that $\langle w, Rw \rangle \rightarrow \infty$ for $\|w\| \rightarrow \infty$ and $w \in \text{Im}(R)$. Thus the functional $w \mapsto \langle w, Rw \rangle$ attains a global minimum on $\{w \in \mathbb{R}^L | L^\top w = m_i\} \cap \text{Im}(R)$. Let v_1, \dots, v_r denote an orthonormal basis of $\text{Ker}(R)$.⁵ We thus search for a solution of

$$w_M^i(L_n) = \arg \min_{\substack{L^\top w = m_i \\ \langle v_1, w \rangle = 0, \dots, \langle v_r, w \rangle = 0}} \langle w, Rw \rangle \quad \text{for } 1 \leq i \leq n.$$

Using Lagrange multipliers, we thus get for $V = (v_1, \dots, v_r)$ that

$$2Rw = L \cdot \lambda + V \cdot \mu \quad (3.17)$$

$$L^\top w = m_i \quad (3.18)$$

$$V^\top \cdot w = 0 \quad (3.19)$$

⁵Note that we have either $\text{Ker}(R) = \{0\}$ in the MEG case or $\text{Ker}(R) = \mathbb{R} \cdot 1$ in the EEG case. This argument, however, works for an arbitrary r .

for Lagrange multipliers $\lambda \in \mathbb{R}^n$ and $\mu \in \mathbb{R}^r$. Multiplying (3.17) by V^\top and remembering that the lead field range is the same as the image of R , we get $\mu = 0$. By (3.19) we furthermore have $w \in \text{Im}(R)$, which together with (3.17) implies $w = \frac{1}{2}R^{-1}L \cdot \lambda$. Inserting this into (3.18) yields

$$m_i = \frac{1}{2}L^\top R^{-1}L \cdot \lambda \iff \lambda = 2 \cdot (L^\top R^{-1}L)^{-1} \cdot m_i.$$

In combination with the previously derived equation for w , we thus arrive at

$$w_M^i(L_n) = R^{-1}L (L^\top R^{-1}L)^{-1} \cdot m_i.$$

Combining these results for $i = 1, \dots, n$ yields the statement of the lemma. \square

By inserting the expression for the beamformer filter (3.15) into the expression for the beamformer power (3.14), we see that the beamformer power is given by

$$P_M(L_n) := \mathbb{E} [\|W_M(L_n)^\top \cdot d\|_2^2] = \text{Tr}(W_M(L_n)^\top R W_M(L_n)) \quad (3.20)$$

$$= \text{Tr}\left(M^\top (L_n^\top R^{-1}L_n)^{-1} M\right) \quad (3.21)$$

$$= \text{Tr}\left((L_n^\top R^{-1}L_n)^{-1} M M^\top\right). \quad (3.22)$$

The value $P_M(L_n)$ is now used as a measure to assess how well the given source configuration can explain the data. To reconstruct neural activity using beamforming, one then first estimates the covariance matrix R , see e.g. [51], and then searches for source configurations with a large value of $P_M(L_n)$.

Definition 3.1.5 (UG, NAI, and SAM beamformers). Let $N \in \mathbb{R}^{L \times L}$ denote the noise covariance matrix of the measurement noise. Similar to the covariance matrix R , we also assume that N is positive definite on the lead field range and has the orthogonal complement of the lead field range as its kernel. The common choices for M are as follows.

1. Choosing $M = \text{Id}$, where Id denotes the identity matrix, is called *unit gain (UG)* beamforming, and corresponds to the power estimate

$$\begin{aligned} P_{\text{UG}}(L_n) &:= P_{\text{Id}}(L_n) \\ &= \text{Tr}\left((L_n^\top R^{-1}L_n)^{-1}\right) \end{aligned}$$

2. Choosing $M = (L_n^\top N^{-1}L_n)^{\frac{1}{2}}$ is called *neural activity index (NAI)* beamforming, and corresponds to the power estimate

$$\begin{aligned} P_{\text{NAI}}(L_n) &:= P_{(L_n^\top N^{-1}L_n)^{\frac{1}{2}}}(L_n) \\ &= \text{Tr}\left((L_n^\top R^{-1}L_n)^{-1} (L_n^\top N^{-1}L_n)\right) \end{aligned}$$

3. Choosing $M = (L_n^\top R^{-1}L_n) \cdot (L_n^\top R^{-1}N R^{-1}L_n)^{-\frac{1}{2}}$ is called *synthetic aperture magnetometry (SAM)* beamforming, and corresponds to the power estimate

$$\begin{aligned} P_{\text{SAM}}(L_n) &:= P_{(L_n^\top R^{-1}L_n) \cdot (L_n^\top R^{-1}N R^{-1}L_n)^{-\frac{1}{2}}}(L_n) \\ &= \text{Tr}\left((L_n^\top R^{-1}L_n) \cdot (L_n^\top R^{-1}N R^{-1}L_n)^{-1}\right). \end{aligned}$$

We have now derived the beamformer equations as they are presented in [135]. While the UG beamformer corresponds at first sight to the most “intuitive” choice for M , the problem with the UG approach is that the lead field norm has a significant impact on the filter norm. The UG side constraint in (3.13) implies $\|W_{\text{UG}}^\top \cdot L\| = 1$. If $\|L\|$ is small, $\|W_{\text{UG}}\|$ thus has to be large. But if we now apply $\|W_{\text{UG}}\|$ to measurement data, we will typically not be able to filter out all of the noise, and thus the remaining noise components get amplified for sources with a small lead field norm. If we then, e.g., compare a source close to the head surface with a source that is deep inside the brain, we have that the superficial source is much closer to the sensors than the deep source, and the deep source will hence produce a lead field with a smaller norm. Due to this, the power of the UG filter output will typically increase when one goes from shallow sources to deep sources, which is the cause for the aforementioned “center of head bias”, as it is, e.g., called in [231]. The choices of M for the NAI and the SAM, on the other hand, take out the dependency on the lead field norm. Note, e.g., that $P_{\text{NAI}}(\lambda \cdot L_n) = P_{\text{NAI}}(L_n)$ and $P_{\text{SAM}}(\lambda \cdot L_n) = P_{\text{SAM}}(L_n)$ for an arbitrary $\lambda \neq 0$. Thus, in contrast to the UG beamformer, the NAI and SAM beamformers do not suffer from a lead field norm-induced amplification or attenuation of the noise component in the data.

Furthermore, a straightforward computation shows that for an arbitrary choice of (by assumption full rank) $M \in \mathbb{R}^{n \times n}$ we have for $W = W_M(L_n)$ that

$$\text{Tr} \left(\left(W^\top R W \right) \cdot \left(W^\top N W \right)^{-1} \right) = P_{\text{SAM}}(L_n).$$

Note that the left-hand side of the above equation has the structure of “projected signal power” divided by “projected noise power”, and is hence typically called pseudo-Z score in the literature. From this perspective, one can thus also interpret a large SAM power as a large SNR in the filtered signal. If we particularly focus on the case where our source configuration consists of a single source, i.e., $k = 1$ and $L_1 = l_{x_1, M_1}$, we see that

$$P_{\text{SAM}}(l_{x_1, M_1}) = \frac{W_M(l_{x_1, M_1})^\top \cdot R \cdot W_M(l_{x_1, M_1})}{W_M(l_{x_1, M_1})^\top \cdot N \cdot W_M(l_{x_1, M_1})}.$$

This expression is the original definition of the pseudo-Z score and the SAM beamformer by its inventors [222], and the way it is typically presented in the literature.

Furthermore, note that for $N = \sigma^2 \cdot \text{Id}$, where Id denotes the identity matrix, and for a source configuration consisting of a single dipole, we have

$$P_{\text{SAM}}(l_{x_1, M_1}) = \frac{1}{\sigma^2} \cdot \frac{l_{x_1, M_1}^\top \cdot R^{-1} \cdot l_{x_1, M_1}}{l_{x_1, M_1}^\top \cdot R^{-2} \cdot l_{x_1, M_1}}.$$

In this special case, the SAM beamformer is also called the (scalar) unit noise gain (UNG) beamformer [190]. Similarly, for $N = \sigma^2 \cdot \text{Id}$ with a source configuration consisting of a single dipole, we also have

$$P_{\text{NAI}}(l_{x_1, M_1}) = \frac{1}{\sigma^2} \cdot \frac{l_{x_1, M_1}^\top \cdot l_{x_1, M_1}}{l_{x_1, M_1}^\top \cdot R^{-1} \cdot l_{x_1, M_1}}.$$

In this special case, the NAI beamformer is also called the (scalar) array gain (AG) beamformer [190].

3.2 Dipole scanning and sLORETA

3.2.1 Identification of sLORETA and dipole scanning

We start by deriving explicit formulas for the metrics used in dipole scanning. Since we will use them later on in the beamforming investigation, we will formulate the statements in a more general form than they are needed for the identification of sLORETA and dipole scanning. Furthermore, we will in the following assume that for each source position x_0 the lead field $L_{x_0} \in \mathbb{R}^{L \times 3}$ has rank 3.

Lemma 3.2.1. *Let $A \in \mathbb{R}^{L \times k}$ with $\text{rank}(A) = k$ and $\text{Im}(A)$ contained in the lead field range. Let C be positive definite on the lead field range. Additionally, let $d \in \mathbb{R}^L$ be contained in the lead field range. Then*

$$\arg \min_{j \in \mathbb{R}^k} \|d - A \cdot j\|_C^2 = (A^\top C A)^{-1} A^\top C d.$$

Proof. Since C is positive definite on the lead field range, we can find a symmetric matrix $C^{\frac{1}{2}} \in \mathbb{R}^{L \times L}$ such that for each v in the lead field range we have $C \cdot v = C^{\frac{1}{2}} \cdot C^{\frac{1}{2}} \cdot v$. Using $\|x\|_C^2 = \langle Cx, x \rangle$, we then have for $j \in \mathbb{R}^k$ that

$$\begin{aligned} \|d - A \cdot j\|_C^2 &= \langle C(d - A \cdot j), d - A \cdot j \rangle = \langle C^{\frac{1}{2}}(d - A \cdot j), C^{\frac{1}{2}}(d - A \cdot j) \rangle \\ &= \|C^{\frac{1}{2}}d - C^{\frac{1}{2}}A \cdot j\|_2^2. \end{aligned}$$

We have thus rewritten the problem as an ordinary least squares problem, whose solution $j_C(d, A) \in \mathbb{R}^k$ is characterized by the Gauss normal equation

$$\left(C^{\frac{1}{2}}A\right)^\top \left(C^{\frac{1}{2}}A\right) \cdot j_C(d, A) = \left(C^{\frac{1}{2}}A\right)^\top \cdot \left(C^{\frac{1}{2}}d\right),$$

see e.g. [203], Theorem 11.1. Since C is positive definite on the lead field range, the matrix $A^\top C A$ is positive definite and hence invertible, and the last equation is equivalent to

$$j_C(d, A) = (A^\top C A)^{-1} A^\top C d.$$

□

Lemma 3.2.2. *Let $A \in \mathbb{R}^{L \times k}$, $C \in \mathbb{R}^{L \times L}$ and $d \in \mathbb{R}^L$ be as in lemma 3.2.1. We then have*

$$\text{rv}_C(d, A) = \|d\|_C^2 - d^\top C A (A^\top C A)^{-1} A^\top C d \quad (3.23)$$

$$\text{GOF}_C(d, A) = \frac{d^\top C A (A^\top C A)^{-1} A^\top C d}{\|d\|_C^2} \quad (3.24)$$

Proof. By lemma 3.2.1, we have

$$j_C(d, A) := \arg \min_{j \in \mathbb{R}^3} \|d - A \cdot j\|_C^2 = (A^\top C A)^{-1} A^\top C d.$$

This implies

$$\begin{aligned} \text{rv}_C(d, A) &= \|d - A \cdot j_C(d, A)\|_C^2 = \|d\|_C^2 - 2\langle d, A \cdot j_C(d, A) \rangle_C + \|A \cdot j_C(d, A)\|_C^2 \\ &= \|d\|_C^2 - 2d^\top C A (A^\top C A)^{-1} A^\top C d + \|A \cdot j_C(d, A)\|_C^2. \end{aligned}$$

We can now compute

$$\begin{aligned}\|A \cdot j_C(d, A)\|_C^2 &= j_C(d, A)^\top A^\top C A j_C(d, A) \\ &= d^\top C A (A^\top C A)^{-1} A^\top C A (A^\top C A)^{-1} A^\top C d \\ &= d^\top C A (A^\top C A)^{-1} A^\top C d,\end{aligned}$$

and thus

$$\text{rv}_C(d, A) = \|d\|_C^2 - d^\top C A (A^\top C A)^{-1} A^\top C d.$$

This, in turn, implies

$$\text{GOF}_C(d, A) = 1 - \frac{\text{rv}_C(d, A)}{\|d\|_C^2} = \frac{d^\top C A (A^\top C A)^{-1} A^\top C d}{\|d\|_C^2}.$$

□

Theorem 3.2.3 (Equivalence of sLORETA and dipole scanning). *Let $C \in \mathbb{R}^{L \times L}$ be positive definite on the lead field range. Denote by $L_{x_0} \in \mathbb{R}^{L \times 3}$ the lead field of a position x_0 . We then have*

$$\|j_C^{\text{sLORETA}}(d, L_{x_0})\|_2^2 = \|d\|_C^2 \cdot \text{GOF}_C(d, L_{x_0}).$$

Hence, we see that the generalized sLORETA estimate at a source position is, up to a lead field independent factor, given by the goodness of fit of the lead field at the investigated source position to the data.

Proof. By definition 3.1.3, we have

$$j_C^{\text{sLORETA}}(d, L_{x_0}) = (L_{x_0}^\top C L_{x_0})^{-\frac{1}{2}} L_{x_0}^\top C d.$$

By lemma 3.2.2, we thus have

$$\begin{aligned}\|j_C^{\text{sLORETA}}(d, L_{x_0})\|_2^2 &= j_C^{\text{sLORETA}}(d, L_{x_0})^\top \cdot j_C^{\text{sLORETA}}(d, L_{x_0}) \\ &= d^\top C L_{x_0} (L_{x_0}^\top C L_{x_0})^{-1} L_{x_0}^\top C d \\ &= \|d\|_C^2 \cdot \text{GOF}_C(d, L_{x_0}).\end{aligned}$$

□

Now note that the scalar $\|d\|_C^2$ in the theorem above is independent of the position x_0 . We thus see that searching for positions with a large value of $\|j_C^{\text{sLORETA}}(d, L_{x_0})\|_2^2$ is equivalent to searching for positions with a large goodness of fit $\text{GOF}_C(d, L_{x_0})$. This means that performing source reconstruction using any member of the family of generalized sLORETA approaches is the same as searching for source positions from where a single dipolar source can well explain the measured data, i.e., sLORETA is a dipole scan.

One particularly interesting consequence of this identification is that it immediately makes clear why sLORETA is able to correctly localize noiseless single dipolar sources, since dipole scans have this property almost by definition.

Theorem 3.2.4. *Assume that the signal vector $d \in \mathbb{R}^L$ originates from a dipole at position x_0 with moment M , i.e. $d = L_{x_0} \cdot M$. Let $C \in \mathbb{R}^{L \times L}$ be positive definite on the lead field range. Then the function*

$$\{A \in \mathbb{R}^{L \times 3} \mid \text{rank}(A) = 3\} \rightarrow \mathbb{R}; A \mapsto \|j_C^{\text{sLORETA}}(d, A)\|_2^2$$

attains a global maximum at $A = L_{x_0}$, i.e., at the lead field corresponding to the true source position.

Proof. Note that by construction $\text{GOF}_C(L_{x_0} \cdot M, A) \leq 1$ for any $A \in \mathbb{R}^{L \times 3}$. Now, we have

$$0 = \min_{j \in \mathbb{R}^3} \|L_{x_0} \cdot M - L_{x_0} \cdot j\|_C^2,$$

and thus $\text{rv}_C(L_{x_0} \cdot M, L_{x_0}) = 0$ and $\text{GOF}_C(L_{x_0} \cdot M, L_{x_0}) = 1$. Since by theorem 3.2.3 we have that $A \mapsto \|j_C^{\text{sLORETA}}(L_{x_0} \cdot M, A)\|_2^2$ is a positive multiple of $A \mapsto \text{GOF}_C(L_{x_0} \cdot M, A)$, the claim of the theorem follows. \square

Another straightforward consequence of this identification is that it immediately yields a strategy for estimating source orientations from the sLORETA estimate. As sLORETA can be interpreted as a dipole scan with respect to an inner product induced by a matrix C , the corresponding dipole scan estimate $j_C(d, L_{x_0})$ of the dipole moment at the position x_0 with respect to the measurement data d is given by

$$\begin{aligned} j_C(d, L_{x_0}) &\stackrel{\text{Lemma 3.2.1}}{=} (L_{x_0}^T C L_{x_0})^{-1} \cdot L_{x_0}^T \cdot C \cdot d \\ &\stackrel{(3.4)}{=} (L_{x_0}^T C L_{x_0})^{-\frac{1}{2}} \cdot j_C^{\text{sLORETA}}(d, L_{x_0}). \end{aligned}$$

If we thus multiply the sLORETA vector $j_C^{\text{sLORETA}}(d, L_{x_0})$ by $(L_{x_0}^T C L_{x_0})^{-\frac{1}{2}}$, we arrive at an estimate of the dipole moment at this position. This is illustrated in figure 3.1.

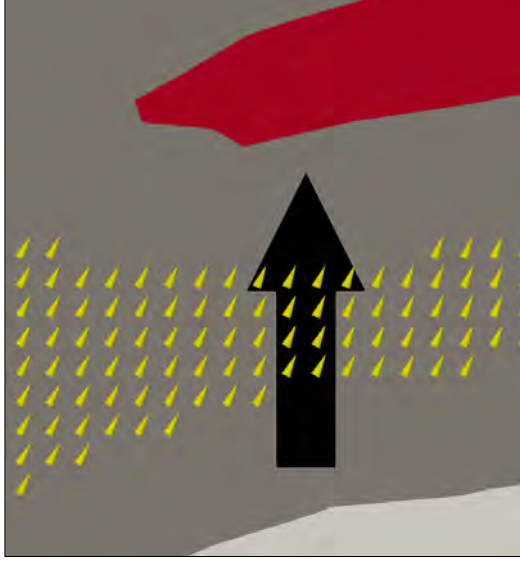
We thus see that if we view sLORETA as a dipole scan, it is straightforward to estimate dipole orientations, while, in the original interpretation as a distributed inverse method, there is no natural way to do this [77]. In our opinion, this is a further strong argument for preferring the dipole scan interpretation of sLORETA over the current density perspective that is ubiquitous in the literature.

3.2.2 The impact of noise

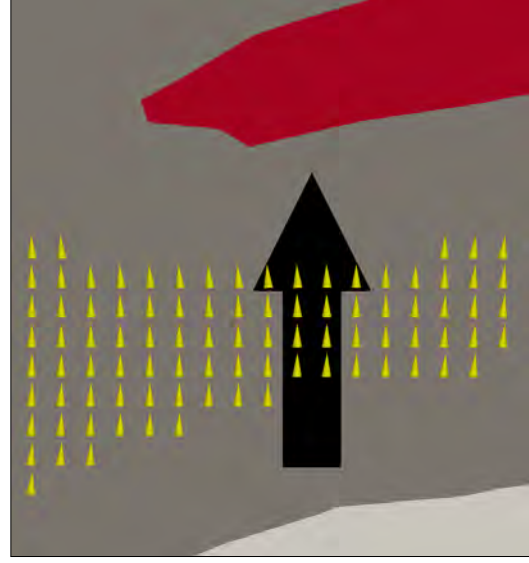
In the paper introducing the original sLORETA approach [149], the author shows empirically that the method is able to correctly localize noiseless dipolar sources, as we have also shown in theorem 3.2.4. And while this is certainly a desirable property, real measurement data always comes with noise. With this in mind, both [192] and [78] almost simultaneously published results showing that, under the assumption of white noise, sLORETA leads to a biased source reconstruction. The author of [149] then later responds to these findings in [150]. There, he models the noise n not as white noise, but as

$$n = n_{\text{meas}} + L^E \cdot n_{\text{bio}}, \quad (3.25)$$

where $\text{cov}(n_{\text{meas}}) = \sigma_{\text{meas}}^2 \cdot p_{\text{CAR}}$, $\text{cov}(n_{\text{bio}}) = \sigma_{\text{bio}}^2 \cdot \text{Id}$, $\sigma_{\text{meas}}, \sigma_{\text{bio}} > 0$, and n_{meas} is independent of n_{bio} . Here, $\text{cov}(X) \in \mathbb{R}^{L \times L}$ denotes the covariance matrix of the random vector X . If one



(a) Raw sLORETA orientations. Note how the reconstructions are tilted with respect to the true orientation



(b) “Corrected” sLORETA orientations. Note how the reconstructed orientations align with the true orientation.

Figure 3.1: Result of an sLORETA reconstruction in a sagittal slice of the realistic volume conductor model from figure 2.1a. In (a), we visualize the vectors $j_C^{\text{sLORETA}}(d, L_{x_0})$ from definition 3.1.3, while in (b) we visualize the vectors $(L_{x_0}^\top C L_{x_0})^{-\frac{1}{2}} \cdot j_C^{\text{sLORETA}}(d, L_{x_0})$. In both (a) and (b), the reconstructions are visualized as yellow cones, while the true dipole is visualized as a black arrow. The source space used in the reconstruction was constructed with a resolution of 0.25 mm and a distance condition of 1 mm from the gray matter boundary. The true dipole is not contained in the source space. For better visibility, we truncated the components of the reconstructions that were not contained in the sagittal plane. The gray matter is visualized in gray, the white matter in white, and the CSF in red.

then chooses $\alpha := \frac{\sigma_{\text{meas}}^2}{\sigma_{\text{bio}}^2}$, it is proven in [150] that the corresponding classical sLORETA method, i.e., the generalized sLORETA approach with the choice

$$C = \left(L^E (L^E)^\top + \alpha \cdot p_{\text{CAR}} \right)^+,$$

as defined in (3.5), yields an unbiased estimation of noisy signals originating from single dipolar sources. The author of [150] then argues that the noise model (3.25) is more appropriate for EEG data than assuming white noise, as (3.25) not only includes additive noise on the sensors, but also background brain activity, which propagates to the measurement data through the lead field. We, in fact, agree that (3.25) is, in the absence of better information, a sensible guess of the noise distribution. Furthermore, the investigation in [150] revealed that the α in (3.5) defining the C matrix of the classical sLORETA method should not be interpreted as some form of regularization parameter, but rather as an estimate of the ratio of sensor noise and biological noise.

Furthermore, the author in [150] shows that the EEG eLORETA approach is unbiased if the

noise n can be described by

$$n = n_{\text{meas}} + L^{\text{E}} \cdot n_{\text{bio}}, \quad (3.26)$$

with $\text{cov}(n_{\text{meas}}) = \sigma_{\text{meas}}^2 \cdot p_{\text{CAR}}$ and $\text{cov}(n_{\text{bio}}) = \sigma_{\text{bio}}^2 \cdot (W^{\text{E}})^{-1}$, where W^{E} is the weight matrix defined by equation (3.9), and one again assumes that the sensor noise is independent of the biological noise. We do not see how the assumption $\text{cov}(n_{\text{bio}}) = \sigma_{\text{bio}}^2 \cdot (W^{\text{E}})^{-1}$ could be justified. And although he does not state it explicitly, from the discussion at the bottom of page 7 in [150] we get the impression that the author himself thinks the same.

Using the perspective that the family of generalized sLORETA approaches can be identified with dipole scans, we can realize these miscellaneous results as special cases of a single general result about the biasedness of dipole scans. The goal of this subsection is the formulation and proof of this theorem.

Definition 3.2.5. For the remainder of this subsection, we assume that the signal d is due to a single dipolar source at position x_0 with moment M and signal strength s , with additive uncorrelated noise n . More concretely, we thus model

$$d = L_{x_0} \cdot M \cdot s + n, \quad (3.27)$$

and we assume $\text{cov}(s, n_i) = 0$ for each component n_i of $n = (n_1, \dots, n_L)$, and $\mathbb{E}[n] = 0$. Furthermore, we define $Q^2 := \mathbb{E}[s^2]$, and we assume $Q^2 > 0$. Then, for each C that is positive definite on the lead field range, the generalized sLORETA estimate

$$j_C^{\text{sLORETA}}(d, L_x) = (L_x^{\text{T}} C L_x)^{-\frac{1}{2}} L_x^{\text{T}} C d$$

from (3.4) is also a random variable. For generalized sLORETA reconstructions, we look for source positions where $\|j_C^{\text{sLORETA}}(d, L_x)\|_2^2$ is large. Accordingly, [150], [192], and [78] define a generalized sLORETA approach to be *unbiased* if, and only if, one has

$$L_{x_0} = \arg \max_{L \text{ lead field}} \mathbb{E} [\|j_C^{\text{sLORETA}}(d, L)\|_2^2], \quad (3.28)$$

i.e., if the expected value of the metric used for source localization is maximized by the true lead field. Now, by theorem 3.2.3 and lemma 3.2.2, the residual variance and the generalized sLORETA metric are related by

$$\|j_C^{\text{sLORETA}}(d, L)\|_2^2 = \|d\|_2^2 - \text{rv}_C(d, L).$$

This in particular shows that the unbiasedness of the generalized sLORETA method for a parameter matrix C is equivalent to the statement

$$L_{x_0} = \arg \min_{L \text{ lead field}} \mathbb{E} [\text{rv}_C(d, L)],$$

i.e., in the dipole scanning interpretation of the generalized sLORETA method, unbiasedness is equivalent to requiring that the expected residual variance of a lead field to the data is minimized in the true lead field.

For the remainder of this subsection, we will only discuss the MEG case, where we assume the noise covariance matrix $N := \text{cov}(n) \in \mathbb{R}^{L \times L}$ to be invertible. By restricting to the subspace $\mathbb{R}_{\text{CAR}}^L$ and adapting all definitions accordingly, the results of this subsection can be directly transferred to the EEG case. But as this would introduce a lot of boilerplate text that, in essence, uses the same ideas as in the MEG case, we will refrain from doing so in this thesis.

Theorem 3.2.6. *Without further assumptions on the lead fields, we can guarantee*

$$L_{x_0} = \arg \min_{L \text{ lead field}} \mathbb{E} [\text{rv}_C(d, L)] \quad (3.29)$$

if, and only if, $C = \alpha \cdot N^{-1}$ for some $\alpha > 0$.

Note that the choice $C = N^{-1}$ corresponds to a pre-whitening of the data. Furthermore, note that performing a dipole scan with $C = \alpha \cdot N^{-1}$ is equivalent to performing a standard L2 dipole scan on pre-whitened data. We will now work towards a proof of theorem 3.2.6.

Lemma 3.2.7. *We have that*

$$\begin{aligned} & \arg \min_{L \text{ lead field}} \mathbb{E} [\text{rv}_C(d, L)] \\ &= \arg \max_{L \text{ lead field}} Q^2 M^\top (L_{x_0}^\top CL) (L^\top CL)^{-1} (L^\top CL_{x_0}) M + \text{Tr} \left((L^\top CL)^{-1} L^\top CNCL \right). \end{aligned}$$

Proof. In definition 3.2.5 we derived

$$\arg \min_{L \text{ lead field}} \mathbb{E} [\text{rv}_C(d, L)] = \arg \max_{L \text{ lead field}} \mathbb{E} [\|j_C^{\text{sLORETA}}(d, L)\|^2],$$

which yields for $L \in \mathbb{R}^{L \times 3}$ that

$$\begin{aligned} \mathbb{E} [\|j_C^{\text{sLORETA}}(d, L)\|^2] &\stackrel{(3.4)}{=} \mathbb{E} \left[d^\top CL (L^\top CL)^{-1} L^\top C d \right] \\ &\stackrel{(3.27)}{=} \mathbb{E} \left[s^2 M^\top (L_{x_0}^\top CL) (L^\top CL)^{-1} (L^\top CL_{x_0}) M + n^\top CL (L^\top CL)^{-1} L^\top C n \right. \\ &\quad \left. + (sn^\top) CL (L^\top CL)^{-1} L^\top CL_{x_0} M + M^\top L_{x_0}^\top CL (L^\top CL)^{-1} L^\top C (sn) \right] \\ &= Q^2 M^\top (L_{x_0}^\top CL) (L^\top CL)^{-1} (L^\top CL_{x_0}) M + \text{Tr} \left(CL (L^\top CL)^{-1} L^\top CN \right) \end{aligned}$$

and thus finishes the proof of the lemma. \square

Now let $d_0 := QL_{x_0}M$. Then

$$\begin{aligned} Q^2 M^\top (L_{x_0}^\top CL) (L^\top CL)^{-1} (L^\top CL_{x_0}) M &= d_0^\top CL (L^\top CL)^{-1} L^\top C d_0 \\ &\stackrel{\text{Lemma 3.2.2}}{=} \|d_0\|_C^2 \cdot \text{GOF}_C(d_0, L), \end{aligned}$$

and we see that the right-hand side of lemma 3.2.7 is essentially given by the sum of the goodness of fit of the lead field L to the noiseless dipolar signal $d_0 = QL_{x_0}\eta$ and a noise-dependent distortion. We can now prove one half of theorem 3.2.6.

Proof that $C = \alpha \cdot N^{-1}$ is sufficient for equation (3.29) to hold. If $C = \alpha \cdot N^{-1}$, we have

$$\text{Tr} \left((L^\top CL)^{-1} L^\top CNCL \right) = 3\alpha,$$

i.e. the noise-dependent distortion in lemma 3.2.7 is independent of the lead field L , and we hence have

$$\arg \min_{L \text{ lead field}} \mathbb{E} [\text{rv}_C(d, L)] = \arg \max_{L \text{ lead field}} \|d_0\|_C^2 \cdot \text{GOF}_C(d_0, L) = L_{x_0}.$$

\square

Before we start to approach the other half of theorem 3.2.6, we first give a quick recapitulation on differentiability in normed spaces.

Definition 3.2.8. Let V, W denote normed vector spaces, let $U \subset V$ be open, and let $f : U \rightarrow W$ be a map. We then say that f is *Fréchet differentiable*, or simply *differentiable*, in $x \in U$ if there exists a bounded linear operator $T : V \rightarrow W$ such that

$$f(x + h) = f(x) + T(h) + \varphi(h),$$

where $\varphi(h) \in o(h)$. Such an operator T , if it exists, is unique. We then write $T = Df(x)$ and call $Df(x)$ the *derivative* of f at x .

The usual properties of derivatives, to the extent that they make sense in general normed spaces, carry over to Fréchet derivatives, see e.g. [44], Chapter VIII. One can, for example, show that the chain rule holds (see [44], (8.2.1)). Furthermore, if a scalar differentiable function f has a local optimum in x , it follows that $Df(x) = 0$ (see [44], Section 8.2, Problem 3). In particular, if $Df(x) \neq 0$ then f cannot have a local maximum at x .

Our strategy for the second half of theorem 3.2.6 is now as follows. Let

$$O = \{L \in \mathbb{R}^{L \times 3} \mid \text{rank}(L) = 3\}.$$

Then $O \subset \mathbb{R}^{L \times 3}$ is open and dense, and the maps

$$f : O \rightarrow \mathbb{R}; L \mapsto Q^2 \eta^\top (L_{x_0}^\top C L) (L^\top C L)^{-1} (L^\top C L_{x_0}) \eta \quad (3.30)$$

and

$$g : O \rightarrow \mathbb{R}; L \mapsto \text{Tr} \left((L^\top C L)^{-1} L^\top C N C L \right) \quad (3.31)$$

are well defined. We will now investigate under what conditions we have $D(f + g)(L_{x_0}) \neq 0$, which by lemma 3.2.7 will lead to a proof of theorem 3.2.6.

First, note that as derived above we have $f(L) = \|d_0\|_C^2 \cdot \text{GOF}_C(d_0, L)$, and hence $f(L)$ has a local maximum at L_{x_0} . Thus we must have $Df(L_{x_0}) = 0$. The derivative of g is computed in the following lemma.

Lemma 3.2.9. *Let $L \in O$. Then for $A \in \mathbb{R}^{L \times 3}$ we have*

$$Dg(L)(A) = 2 \cdot \text{Tr} \left(A \cdot \left((L^\top C L)^{-1} L^\top C N C - (L^\top C L)^{-1} L^\top C N C L (L^\top C L)^{-1} L^\top C \right) \right). \quad (3.32)$$

Proof. First note that if $h : O \rightarrow \mathbb{R}^{3 \times 3}$ is differentiable, the map

$$H : O \rightarrow \mathbb{R}^{L \times L}; L \mapsto L \cdot h(L) \cdot L^\top. \quad (3.33)$$

is also differentiable with

$$DH(L)(A) = L \cdot h(L) \cdot A^\top + A \cdot h(L) \cdot L^\top + L \cdot Dh(L)(A) \cdot L^\top. \quad (3.34)$$

This equation can be derived by simply writing out $H(L + A)$ and gathering all terms in $o(A)$. Furthermore, if $U \subset \mathbb{R}^{3 \times 3}$ is the open set of invertible matrices, the map $I : U \rightarrow U; T \mapsto T^{-1}$ is differentiable with

$$DI(T)(S) = -T^{-1} \cdot S \cdot T^{-1}, \quad (3.35)$$

see e.g. [44], (8.3.2). Alternatively, this equation can be derived by a direct computation using the Woodbury matrix identity.⁶ If we now define

$$g_1 : O \rightarrow U; L \mapsto L^\top CL,$$

it is an easy exercise to see that

$$Dg_1(L)(A) = L^\top CA + A^\top CL. \quad (3.36)$$

If we thus define $h := I \circ g_1$, and H as in (3.33), we see that

$$H(L) = L \cdot (L^\top CL)^{-1} \cdot L^\top,$$

and the combination of (3.34), (3.35), (3.36), and the chain rule yield that

$$DH(L)(A) = L \cdot (L^\top CL)^{-1} \cdot A^\top + A \cdot (L^\top CL)^{-1} \cdot L^\top - L \cdot (L^\top CL)^{-1} \cdot (L^\top CA + A^\top CL) \cdot (L^\top CL)^{-1} \cdot L^\top \quad (3.37)$$

Furthermore, it is easy to see that for an arbitrary bounded linear map $T : V \rightarrow W$, one always has $DT(X) = T$ for all $X \in V$, see e.g. [44], (8.1.3). Now, the map

$$T : \mathbb{R}^{L \times L} \rightarrow \mathbb{R}; D \mapsto \text{Tr} \left(N^{\frac{1}{2}} CDCN^{\frac{1}{2}} \right) \quad (3.38)$$

is a bounded linear map in D , and we have $g = T \circ H$, which follows from the property of the trace function that $\text{Tr}(X \cdot Y) = \text{Tr}(Y \cdot X)$ for arbitrary $X \in \mathbb{R}^{i \times j}, Y \in \mathbb{R}^{j \times i}$. But then, the chain rule implies that

$$Dg(L)(A) = T(DH(L)(A)) = \text{Tr} \left(N^{\frac{1}{2}} C \cdot DH(L)(A) \cdot CN^{\frac{1}{2}} \right).$$

Inserting (3.37) into this last expression, followed by a couple of straightforward algebraic manipulations, yields the statement of this lemma. \square

We can now exactly characterize under what conditions we have $Dg(L) = 0$.

Lemma 3.2.10. *The following are equivalent.*

1. $Dg(L) = 0$
2. $C \cdot N \cdot \text{Ker}(L^\top) = \text{Ker}(L^\top)$

Proof. Assume $Dg(L) = 0$. First note that $\text{Tr}(B \cdot A) = 0$ for all $B \in \mathbb{R}^{N \times 3}$ implies $A = 0$, as can be seen from taking $B = A^\top$ and remembering that the trace of a positive semidefinite matrix is the sum of its (non-negative) eigenvalues. Hence, equation (3.32) implies

$$L^\top CN \left(E_N - CL(L^\top CL)^{-1} L^\top \right) = 0$$

Hence, for $v \in \text{Ker}(L^\top)$ we have $L^\top CNv = 0$, and hence $C \cdot N \cdot \text{Ker}(L^\top) \subset \text{Ker}(L^\top)$. Since $\dim(C \cdot N \cdot \text{Ker}(L^\top)) = \dim(\text{Ker}(L^\top))$, we must in fact have equality.

On the other hand, assume that we have $C \cdot N \cdot \text{Ker}(L^\top) = \text{Ker}(L^\top)$. Then for $v \in \text{Ker}(L^\top)$ we have

$$L^\top CN \left(E_N - CL(L^\top CL)^{-1} L^\top \right) \cdot v = 0.$$

⁶https://en.wikipedia.org/wiki/Woodbury_matrix_identity

If we now have $v \in \text{Im}(CL)$, we can write $v = CLz$. But then

$$L^\top CN \left(E_N - CL (L^\top CL)^{-1} L^\top \right) \cdot v = L^\top CN (CLz - CLz) = 0.$$

Since $\mathbb{R}^N = \text{Ker}(L^\top) \oplus \text{Im}(CL)$, this implies

$$L^\top CN \left(E_N - CL (L^\top CL)^{-1} L^\top \right) = 0,$$

which by equation (3.32) implies $Dg(L) = 0$. \square

Remark 3.2.11. In the special case of $C = \text{Id}_L$, we have

$$g(L) = \text{Tr} \left((L^\top L)^{-1} L^\top N L \right),$$

and since $\text{Ker}(L^\top)$ is the orthogonal complement with respect to the standard inner product of $\text{Im}(L)$ and N is positive definite, we see that, in this special case, lemma 3.2.10 can equivalently be stated as the equivalence of the following two statements.

1. $Dg(L) = 0$
2. $N \cdot \text{Im}(L) = \text{Im}(L)$

Hence, the differential of g vanishes in the lead field L if, and only if, the image of L is an invariant subspace of N . We thus see that lemma 3.2.10 is a generalization of the well-known fact that, for a positive definite matrix, the gradient of its Rayleigh quotient vanishes exactly in the eigenvectors of the matrix.

Proof of the second half of theorem 3.2.6. By lemma 3.2.7, lemma 3.2.10, and $D(f)(L_{x_0}) = 0$, it follows that we have $D(f + g)(L_{x_0}) = 0$ if, and only if

$$C \cdot N \cdot \text{Ker}(L_{x_0}^\top) = \text{Ker}(L_{x_0}^\top),$$

i.e. if the kernel of $L_{x_0}^\top$ is an invariant subspace of $C \cdot N$. If C is not of the form $\alpha \cdot N^{-1}$ for an $\alpha > 0$, there exists a $(n - 3)$ -dimensional subspace of \mathbb{R}^L that is not invariant under $C \cdot N$. If we make no further assumption on L_{x_0} , we cannot rule out that the kernel of $L_{x_0}^\top$ is such a subspace. We hence cannot guarantee that $C \cdot N \cdot \text{Ker}(L_{x_0}^\top) = \text{Ker}(L_{x_0}^\top)$, and hence we cannot guarantee that $D(f + g)(L_{x_0}) = 0$. But if $D(f + g)(L_{x_0}) \neq 0$, there is some $\tilde{L} \neq L_{x_0}$ such that $(f + g)(\tilde{L}) > (f + g)(L_{x_0})$, and hence

$$L_{x_0} \neq \arg \min_{L \text{ lead field}} \mathbb{E}[\text{rv}_C(d, L)].$$

\square

Remark 3.2.12. We can now realize the results from [192], [78], and [150] as special cases of theorem 3.2.6.

In the case of white noise, i.e., $N = \sigma^2 \cdot \text{Id}_L$, an unbiased generalized sLORETA reconstruction is only possible for the choice $C = \alpha \cdot \text{Id}_L$, i.e., for a standard L2 dipole scan. As both the classical sLORETA approach as well as the eLORETA approach choose a different C , see (3.6) and (3.8), it follows that both the classical sLORETA approach as well as the eLORETA approach are

biased in the presence of white noise. For the classical sLORETA approach in the scalar case, this was derived both in [192] as well as in [78].

If the noise is, however, given by

$$n = n_{\text{meas}} + L \cdot n_{\text{bio}},$$

where $\text{cov}(n_{\text{meas}}) = \sigma_{\text{meas}}^2 \cdot \text{Id}_L$, $\text{cov}(n_{\text{bio}}) = \sigma_{\text{bio}}^2 \cdot \text{Id}$, and the sensor and biological noise are independent, an unbiased reconstruction requires C to be a multiple of the inverse of

$$\text{cov}(N) = \sigma_{\text{meas}}^2 \cdot \text{Id}_L + \sigma_{\text{bio}}^2 \cdot L \cdot L^\top,$$

i.e. C has to be a multiple of $\left(L \cdot L^\top + \frac{\sigma_{\text{meas}}^2}{\sigma_{\text{bio}}^2} \cdot \text{Id}\right)^{-1}$, which corresponds exactly to the classical sLORETA approach as defined in (3.6). The fact that for such a noise model the classical sLORETA approach is unbiased was, to the best of our knowledge, first derived in [150].

Furthermore, if the noise is given by

$$n = n_{\text{meas}} + L \cdot n_{\text{bio}},$$

where $\text{cov}(n_{\text{meas}}) = \sigma_{\text{meas}}^2 \cdot \text{Id}_L$, $\text{cov}(n_{\text{bio}}) = \sigma_{\text{bio}}^2 \cdot (W^M)^{-1}$, and the sensor and biological noise are independent, then the noise covariance matrix is given by

$$\text{cov}(N) = \sigma_{\text{meas}}^2 \cdot \text{Id}_L + \sigma_{\text{bio}}^2 \cdot L \cdot (W^M)^{-1} \cdot L^\top,$$

i.e., for an unbiased estimation, one has to choose

$$C = \rho \cdot \left(L \cdot (W^M)^{-1} \cdot L^\top + \frac{\sigma_{\text{meas}}^2}{\sigma_{\text{bio}}^2} \cdot \text{Id}_L \right)^{-1},$$

for some $\rho > 0$, which corresponds exactly to the eLORETA approach from equation (3.8). The unbiasedness of eLORETA for such a noise model was also derived in [150].

Additionally, note that from the perspective of investigating under what conditions the generalized sLORETA approach is unbiased, we arrive at the conclusion that the α parameter in the definitions of the sLORETA and eLORETA approaches (see (3.5), (3.6), (3.7), (3.8)) should not be interpreted as some form of regularization parameter, but instead as an estimate of the ratio of sensor noise to biological noise.

We have thus seen that the previously published results about the unbiasedness of different versions of the generalized sLORETA approach can be derived as straightforward corollaries of theorem 3.2.6. Moreover, we see that the discussion about whether sLORETA, or eLORETA, or any method of the generalized sLORETA reconstruction family, is unbiased in each case comes down to whether the choice of C matches the noise covariance matrix N . Note that we cannot choose N , while we can choose C . In our opinion, the takeaway from this section should thus be that if one aims for an unbiased reconstruction, one should estimate the noise covariance N and then choose C accordingly, instead of first picking a method and then wondering if the noise fits this choice.⁷

This finishes our discussion on the generalized sLORETA approach. For the remainder of this chapter, we will focus on beamforming approaches.

⁷ N can e.g. be estimated from an empty room measurement or a resting state measurement. We refer to [51] for strategies on how to perform the estimation of N .

3.3 Beamforming investigations

In this section, we want to discuss two topics. First, we saw in the preceding chapter that sLORETA is exactly the same as a dipole scan. We want to present a similar result for beamforming approaches. Concretely, while we cannot identify dipole scanning and beamforming in general, we will show that in the special case that the measurement data can be described by a noisy single dipole, as, e.g., assumed in definition 3.2.5, beamforming is equivalent to performing a dipole scan as defined in section 3.1.2 with the choice $C = N^{-1}$. This implies that if the data can be adequately modeled as originating from a single dipolar source, which is, e.g., oftentimes assumed when trying to reconstruct sources in focal epilepsy [164] or evoked activity [105], then, on a theoretical level, dipole scanning, generalized sLORETA, and beamforming are all equivalent.⁸ Secondly, we want to discuss the feasibility of different beamformer algorithms for reconstructing the orientations of neural sources, which was the topic of our publication [30]. Concretely, we will see that the NAI and SAM beamformers are able to reconstruct the orientations of neural sources, and that, by restricting the power maximization to the tangential plane, they are also able to reconstruct only the tangential components of the orientations. Furthermore, we will investigate the bias of the UG beamformer and will show that, in high SNR scenarios, the UG beamformer can be expected to produce an accurate orientation estimate. We want to highlight that the proofs in this thesis are quite different from their counterparts in [30], as the identification of beamforming and dipole scanning in single-source scenarios enabled a considerable simplification, and an accompanying greater conceptual clarity, of the arguments.

3.3.1 Beamforming and dipole scanning

Throughout this subsection, we will again assume that the data d can be described as a noisy measurement of a single dipolar source at position x_0 with moment M , i.e.

$$d = L_{x_0} \cdot M \cdot s + n, \quad (3.39)$$

where $\text{cov}(s, n_i) = 0$ for each component n_i of $n = (n_1, \dots, n_L)$, $\mathbb{E}[n] = 0$, and $Q^2 := \mathbb{E}[s^2] > 0$. In section 3.1.4 we defined the covariance matrix via $R := \mathbb{E}[d \cdot d^\top] \in \mathbb{R}^{L \times L}$. A straightforward computation then shows that equation (3.39) implies

$$R = N + d_0 \cdot d_0^\top, \quad (3.40)$$

where $N = \mathbb{E}[n \cdot n^\top]$ is the noise covariance matrix and $d_0 = Q \cdot L_{x_0} \cdot M \in \mathbb{R}^L$ is the noiseless dipolar signal. In this subsection, we will again assume that N is invertible. Note, however, that all results are, again, also valid in the common average referenced EEG case, by restricting to the subspace $\mathcal{R}_{\text{CAR}}^L$.

Notice that, if R and N are known, one can reconstruct the linear space $\mathbb{R} \cdot d_0$. One can, e.g., compute

$$N^{-\frac{1}{2}} R N^{-\frac{1}{2}} = \text{Id} + \left(N^{-\frac{1}{2}} d_0 \right) \cdot \left(N^{-\frac{1}{2}} d_0 \right)^\top,$$

and hence $\mathbb{R} \cdot N^{-\frac{1}{2}} d_0$ can be computed as the eigenspace of $N^{-\frac{1}{2}} R N^{-\frac{1}{2}}$ corresponding to the largest (necessarily positive) eigenvalue. From $\mathbb{R} \cdot N^{-\frac{1}{2}} d_0$ we can then compute $\mathbb{R} \cdot d_0$ by multiplying a representative with $N^{\frac{1}{2}}$. If we thus know R and N , we also know $\mathbb{R} \cdot d_0$, and hence also $\text{GOF}_C(d_0, L)$ for an arbitrary positive definite matrix C and lead field matrix $L \in \mathbb{R}^{L \times k}$.⁹

⁸More specifically, the NAI and SAM beamformers are equivalent to dipole scanning. In the special case of white noise, NAI reduces to AG beamforming and SAM reduces to UNG beamforming. The UG beamformer is not equivalent to dipole scanning, but is also, even in the absence of noise, not able to reconstruct single dipoles [190].

⁹Note that $\text{GOF}_C(\lambda \cdot d_0, A) = \text{GOF}_C(d_0, A)$ for all $\lambda \neq 0$, as follows, e.g., from lemma 3.2.2.

A natural approach for source reconstruction is now given as follows.

1. Given R and N , compute an arbitrary vector $0 \neq v \in \mathbb{R} \cdot d_0$.
2. Using this vector, compute $\text{GOF}_C(d_0, L)$ for a set of candidate lead fields L .
3. Search for positions where this value is large.

It is obvious that this provides an unbiased reconstruction of the underlying source, and the maximum value of 1 is attained if, and only if, d_0 is contained in the range of the candidate lead field L . Note that this procedure is just a dipole scan, as it has been defined in section 3.1.2. Nevertheless, the MEEG community has, for whatever reason, decided to call this approach MUSIC [136].¹⁰ To be specific, the common convention in the MEEG community is to speak of a *dipole scan* if the topography d used in the computation of $\text{GOF}_C(d, L)$ arises from a high SNR sample, or from the average of time-locked samples. If the topography d for the computation of $\text{GOF}_C(d, L)$ instead arises from an eigendecomposition of a matrix computed from the noise covariance matrix N and the signal covariance matrix R , it is common to speak of a *MUSIC scan*. We think that this naming convention obfuscates the fact that these two approaches only differ in the step of deriving the signal topography, and will hence not follow it in this thesis. If you do not agree with this choice, you can substitute each occurrence of “dipole scan” in this subsection with “MUSIC scan with a 1-dimensional signal space”.

Our claim is now that SAM beamforming and NAI beamforming, as they are defined in 3.1.5, are equivalent to dipole scanning, both in the scalar and the vector case. As a corollary, this also implies that SAM and NAI beamforming are equivalent to each other in a noisy single-source setting.

The formal statement goes as follows.

Theorem 3.3.1. *There are strictly monotonically increasing functions $f_{d_0, N}, g_{d_0, N} : [0, 1] \rightarrow \mathbb{R}$ such that for all full rank $L \in \mathbb{R}^{n \times k}$ we have*

$$P_{NAI}(L) = k + f_{d_0, N}(\text{GOF}_{N^{-1}}(d_0, L)) \quad (3.41)$$

and

$$P_{SAM}(L) = k + g_{d_0, N}(\text{GOF}_{N^{-1}}(d_0, L)) \quad (3.42)$$

Concretely, these functions are given by

$$f_{d_0, N}(t) = \frac{1}{1 - \mu_{d_0, N} \cdot t} - 1$$

¹⁰We want to note that the general MUSIC algorithm allows signals that are generated by more than a single dipole. If one, however, assumes that the data is generated by a noisy single dipolar source, it is an easy exercise to show that both the scalar MUSIC localizer, as it is for example defined in [136], equation (3), as well as the vector MUSIC localizer, as, e.g., defined in [136], equation (4), are equal to the GOF as it has been calculated in equation (3.24) of this thesis. To be fully explicit, we have the following. Let x be a source position candidate in the brain. Let μ_{vec} denote the vector MUSIC localizer, and μ_s the scalar MUSIC localizer, as they are defined in [136], and let d denote a data vector.

1. Let $L_x \in \mathbb{R}^{L \times 3}$ denote the lead field at position x . For the vector localizer, we have

$$\mu_{\text{vec}}(x) = \text{GOF}_{\text{Id}}(d, L_x).$$

2. If we assume that the orientation η_x of the neural activity at position x is known, the lead field can be reduced to this direction, i.e., $l_x := L_x \cdot \eta_x \in \mathbb{R}^{L \times 1}$. We then have for the scalar localizer

$$\mu_s(x) = \text{GOF}_{\text{Id}}(d, l_x).$$

and

$$g_{d_0, N}(t) = \frac{1}{\langle N^{-1}d_0, d_0 \rangle + 2} \cdot \left(\frac{1}{1 - \rho_{d_0, N} \cdot t} - 1 \right),$$

where the constants are given by

$$\mu_{d_0, N} = \frac{\langle N^{-1}d_0, d_0 \rangle}{1 + \langle N^{-1}d_0, d_0 \rangle}$$

and

$$\rho_{d_0, N} = \frac{\langle N^{-1}d_0, d_0 \rangle^2 + 2\langle N^{-1}d_0, d_0 \rangle}{(\langle N^{-1}d_0, d_0 \rangle + 1)^2}.$$

Note that $0 < \mu_{d_0, N}, \rho_{d_0, N} < 1$.

Remark 3.3.2. Note that this immediately implies that in a noisy single-source setting, both NAI as well as SAM produce an unbiased estimation of the source, i.e., they are maximized by lead fields L whose range contains the true lead field vector. Furthermore, this shows that there is a monotonically increasing function h such that $P_{\text{SAM}}(L) = h(P_{\text{NAI}}(L))$, and we see that the SAM pseudo-Z score and the neural activity index can be transformed into one another via a bijective transformation. They thus contain the same amount of information.

We will now work towards a proof of theorem 3.3.1. During the following computations, we will make use of the equation

$$\text{GOF}_{N^{-1}}(d_0, L) = \frac{d_0^\top N^{-1} L (L^\top N^{-1} L)^{-1} L^\top N^{-1} d_0}{\langle N^{-1}d_0, d_0 \rangle}$$

from lemma 3.2.2 multiple times.

Lemma 3.3.3. *We use the definitions from theorem (3.3.1). Then*

$$\begin{aligned} R^{-1} &= N^{-1} - \frac{\mu_{d_0, N}}{\langle N^{-1}d_0, d_0 \rangle} (N^{-1}d_0) (N^{-1}d_0)^\top \\ R^{-1}NR^{-1} &= N^{-1} - \frac{\rho_{d_0, N}}{\langle N^{-1}d_0, d_0 \rangle} (N^{-1}d_0) (N^{-1}d_0)^\top \end{aligned}$$

Proof. Since $R = N + d_0 d_0^\top$, this follows from a straightforward computation using the Sherman-Morrison formula.¹¹ \square

Proof of Theorem 3.3.1. We start with the neural activity index. Using lemma 3.3.3, we have for an arbitrary full rank $L \in \mathbb{R}^{L \times k}$

$$L^\top R^{-1} L = L^\top N^{-1} L - \frac{\mu_{d_0, N}}{\langle N^{-1}d_0, d_0 \rangle} (L^\top N^{-1} d_0) (L^\top N^{-1} d_0)^\top, \quad (3.43)$$

and thus

$$L^\top N^{-1} L = L^\top R^{-1} L + \frac{\mu_{d_0, N}}{\langle N^{-1}d_0, d_0 \rangle} (L^\top N^{-1} d_0) (L^\top N^{-1} d_0)^\top.$$

This implies

$$(L^\top N^{-1} L) \cdot (L^\top R^{-1} L)^{-1} = \text{Id}_k + \frac{\mu_{d_0, N}}{\langle N^{-1}d_0, d_0 \rangle} (L^\top N^{-1} d_0) (L^\top N^{-1} d_0)^\top \cdot (L^\top R^{-1} L)^{-1}.$$

¹¹https://en.wikipedia.org/wiki/Sherman%E2%80%93Morrison_formula

Using $\text{Tr}(X \cdot Y) = \text{Tr}(Y \cdot X)$, we then have

$$\begin{aligned} P_{\text{NAI}}(L) &= \text{Tr} \left((L^\top N^{-1} L) \cdot (L^\top R^{-1} L)^{-1} \right) \\ &= k + \frac{\mu_{d_0, N}}{\langle N^{-1} d_0, d_0 \rangle} \cdot (L^\top N^{-1} d_0)^\top \cdot (L^\top R^{-1} L)^{-1} \cdot L^\top N^{-1} d_0. \end{aligned}$$

Applying the Sherman-Morrison formula to (3.43), we have

$$(L^\top R^{-1} L)^{-1} = (L^\top N^{-1} L)^{-1} + \frac{\mu_{d_0, N}}{\langle N^{-1} d_0, d_0 \rangle} \cdot \frac{(L^\top N^{-1} L)^{-1} L^\top N^{-1} d_0 (L^\top N^{-1} d_0)^\top (L^\top N^{-1} L)^{-1}}{1 - \mu_{d_0, N} \cdot \text{GOF}_{N^{-1}}(d_0, L)}$$

Putting the last two equations together, we get

$$\begin{aligned} P_{\text{NAI}}(L) - k &= \mu_{d_0, N} \cdot \text{GOF}_{N^{-1}}(d_0, L) + \mu_{d_0, N}^2 \cdot \frac{\text{GOF}_{N^{-1}}(d_0, L)^2}{1 - \mu_{d_0, N} \cdot \text{GOF}_{N^{-1}}(d_0, L)} \\ &= \mu_{d_0, N} \cdot \text{GOF}_{N^{-1}}(d_0, L) \cdot \left(1 + \frac{\mu_{d_0, N} \cdot \text{GOF}_{N^{-1}}(d_0, L)}{1 - \mu_{d_0, N} \cdot \text{GOF}_{N^{-1}}(d_0, L)} \right) \\ &= \frac{\mu_{d_0, N} \cdot \text{GOF}_{N^{-1}}(d_0, L)}{1 - \mu_{d_0, N} \cdot \text{GOF}_{N^{-1}}(d_0, L)} = f_{d_0, N}(\text{GOF}_{N^{-1}}(d_0, L)). \end{aligned}$$

Continuing with the SAM beamformer, we have by lemma 3.3.3 that

$$L^\top R^{-1} N R^{-1} L = L^\top N^{-1} L - \frac{\rho_{d_0, N}}{\langle N^{-1} d_0, d_0 \rangle} (L^\top N^{-1} d_0) (L^\top N^{-1} d_0)^\top \quad (3.44)$$

and

$$L^\top R^{-1} L = L^\top N^{-1} L - \frac{\mu_{d_0, N}}{\langle N^{-1} d_0, d_0 \rangle} (L^\top N^{-1} d_0) (L^\top N^{-1} d_0)^\top,$$

and thus

$$L^\top R^{-1} L = L^\top R^{-1} N R^{-1} L + \frac{\rho_{d_0, N} - \mu_{d_0, N}}{\langle N^{-1} d_0, d_0 \rangle} \cdot (L^\top N^{-1} d_0) (L^\top N^{-1} d_0)^\top.$$

We can now compute

$$\begin{aligned} P_{\text{SAM}}(L) &= \text{Tr} \left((L^\top R^{-1} L) (L^\top R^{-1} N R^{-1} L)^{-1} \right) \\ &= k + \frac{\rho_{d_0, N} - \mu_{d_0, N}}{\langle N^{-1} d_0, d_0 \rangle} \cdot (L^\top N^{-1} d_0)^\top (L^\top R^{-1} N R^{-1} L)^{-1} L^\top N^{-1} d_0. \end{aligned}$$

Applying the Sherman-Morrison formula to equation (3.44), we arrive at

$$\begin{aligned} &(L^\top R^{-1} N R^{-1} L)^{-1} \\ &= (L^\top N^{-1} L)^{-1} + \frac{\rho_{d_0, N}}{\langle N^{-1} d_0, d_0 \rangle} \cdot \frac{(L^\top N^{-1} L)^{-1} (L^\top N^{-1} d_0) (L^\top N^{-1} d_0)^\top (L^\top N^{-1} L)^{-1}}{1 - \rho_{d_0, N} \cdot \text{GOF}_{N^{-1}}(d_0, L)}, \end{aligned}$$

and thus

$$\begin{aligned}
P_{\text{SAM}}(L) - k &= (\rho_{d_0,N} - \mu_{d_0,N}) \cdot \text{GOF}_{N-1}(d_0, L) \\
&\quad + \frac{\rho_{d_0,N} - \mu_{d_0,N}}{\langle N^{-1}d_0, d_0 \rangle} \cdot \frac{\rho_{d_0,N}}{\langle N^{-1}d_0, d_0 \rangle} \cdot \frac{(\langle N^{-1}d_0, d_0 \rangle \cdot \text{GOF}_{N-1}(d_0, L))^2}{1 - \rho_{d_0,N} \cdot \text{GOF}_{N-1}(d_0, L)} \\
&= (\rho_{d_0,N} - \mu_{d_0,N}) \cdot \text{GOF}_{N-1}(d_0, L) \cdot \left(1 + \frac{\rho_{d_0,N} \cdot \text{GOF}_{N-1}(d_0, L)}{1 - \rho_{d_0,N} \cdot \text{GOF}_{N-1}(d_0, L)} \right) \\
&= (\rho_{d_0,N} - \mu_{d_0,N}) \cdot \frac{\text{GOF}_{N-1}(d_0, L)}{1 - \rho_{d_0,N} \cdot \text{GOF}_{N-1}(d_0, L)} \\
&= \frac{\rho_{d_0,N} - \mu_{d_0,N}}{\rho_{d_0,N}} \cdot \left(\frac{1}{1 - \rho_{d_0,N} \cdot \text{GOF}_{N-1}(d_0, L)} - 1 \right) \\
&= \frac{1}{\langle N^{-1}d_0, d_0 \rangle + 2} \left(\frac{1}{1 - \rho_{d_0,N} \cdot \text{GOF}_{N-1}(d_0, L)} - 1 \right) = g_{d_0,N}(\text{GOF}_{N-1}(d_0, L)).
\end{aligned}$$

□

Note that for $0 < \gamma < 1$ we have that for $t \in (0, 1)$

$$\left(\frac{1}{1 - \gamma \cdot t} \right)' = \frac{\gamma}{(1 - \gamma \cdot t)^2} > 0,$$

which shows that $f_{d_0,N}$ and $g_{d_0,N}$ from theorem 3.3.1 are indeed strictly monotonically increasing on $[0, 1]$.

As an application of theorem 3.3.1, we want to discuss so-called *vector beamformers*. In the MEEG community, one speaks of a vector beamformer if the filter W as described in definition 3.13 is of the form $W \in \mathbb{R}^{L \times 3}$. Note that this is the case as long as the lead field L used in the definition has the form $L \in \mathbb{R}^{L \times 3}$. Such a vector beamformer typically arises when investigating the neural activity at some specific position in the brain, say x . If the orientation of the neural activity at position x is not known, or if one does not wish to restrict the orientation to one direction, it is a natural idea to allow all signals originating from position x to pass through the filter. Accordingly, one can choose a source configuration consisting of three sources, each at position x and with the unit directions as moments, for the computation of the beamformer as given in equation (3.13), i.e., one chooses $L_x = (l_{x,(1,0,0)}, l_{x,(0,1,0)}, l_{x,(0,0,1)}) \in \mathbb{R}^{L \times 3}$, see also footnote 2. As discussed in the introduction in section 3.1.4, one can then interpret the filter output of this beamformer as a measure of the neural activity at position x .

As mentioned in the introduction of this chapter, beamforming-based source estimation proceeds by searching for source configurations where the power of the filter output is large. In the vector beamformer setting described in the preceding paragraph, source configurations are chosen to consist of three dipoles at a single position, and the vector beamformer thus measures the source activity at the given position. One can then try to estimate the location of neural activity by iterating over each position in a source space, computing the associated vector beamformer power, and looking for positions where this value is large. We now want to investigate whether this strategy, applied to the different vector beamformers that are common in the MEEG community, is able to correctly localize a noisy single dipolar source.

Definition 3.3.4. Let $O := \{L \in \mathbb{R}^{L \times 3} \mid \text{rank}(L) = 3\}$. We call a function

$$P : O \rightarrow \mathbb{R}; L \mapsto P(L)$$

a *localizer function* if source localization is performed by searching for positions x in the brain where $P(L_x)$ is large. Here $L_x \in \mathbb{R}^{L \times 3}$ denotes the lead field at position x .¹²

Now assume a noisy single-source scenario, i.e., the data d is modelled by

$$d = L_{x_0} \cdot M \cdot s + n,$$

for an arbitrary $L_{x_0} \in \mathbb{R}^{L \times 3}$, $M \in \mathbb{R}^3$, $\text{cov}(s, n_i) = 0$ for each component n_i of $n = (n_1, \dots, n_L)$, and $\mathbb{E}[n] = 0$. We call a strategy for constructing localizer functions *unbiased* if for each noisy single-source scenario the corresponding localizer function attains a global maximum in the true lead field, and *biased* otherwise.

Plainly speaking, we thus say that a reconstruction strategy is unbiased if it is able to correctly reconstruct single dipoles. This, of course, makes no claim about the performance of the method in multiple source scenarios. We, however, view this as a minimum condition for an approach to be viable for the reconstruction of low-rank signals.

In the context of beamforming, the localizer functions are derived from the signal covariance matrix $R = \mathbb{E}[d \cdot d^\top]$. For vector beamformers, there are, roughly speaking, three different classes of localizer function construction strategies in the literature.

1. Probably the most intuitive strategy is to take the power of the post filter signal as a localizer function. Concretely, a vector beamforming approach, in essence, consists of a choice of weighting matrix $M(L) \in \mathbb{R}^{3 \times 3}$. This in turn yields a filter $W(L)$ as defined in (3.13), which, when applied to the measurement data, results in the power

$$P_{M(L)}(L) \stackrel{(3.21)}{=} \text{Tr} \left(M(L)^\top \cdot (L^\top R^{-1} L)^{-1} \cdot M(L) \right)$$

In [190], this is called *power estimate of the first kind* (see eq. (2.77) in that book).

2. Under the assumption that the true source is a single dipole, one might first try to extract the orientation of this dipole from the data, and then use the filter power in this direction as a localizer function. More concretely, if $W(L)$ is a beamformer filter for the lead field L as defined in (3.13), one can define

$$\eta_{\text{opt}}(L) = \arg \max_{\|\eta\|=1} \mathbb{E}[\langle \eta, W(L)^\top d \rangle^2]$$

and then use the map

$$L \mapsto P_{M(L)}^{\eta_{\text{opt}}(L)}(L) := \mathbb{E}[\langle \eta_{\text{opt}}(L), W(L)^\top d \rangle^2]$$

as a localizer function. In [190], this is called *power estimate of the second kind* (see eq. (2.79) in [190]). It is an easy exercise to see that $P_{M(L)}^{\eta_{\text{opt}}(L)}(L)$ can be computed as the largest (necessarily positive) eigenvalue of the matrix

$$M(L)^\top \cdot (L^\top R^{-1} L)^{-1} \cdot M(L).$$

3. In [210], the authors define the so-called *neural activity index* (NAI) as

$$\text{NAI}(L) := \frac{\text{Tr} \left((L^\top R^{-1} L)^{-1} \right)}{\text{Tr} \left((L^\top N^{-1} L)^{-1} \right)}, \quad (3.45)$$

¹²Note that the localizer function is constructed based on the measurement data d . P might, e.g., be the GOF of a lead field to the data vector, or a beamformer power.

and propose to use it as a localizer function. We want to emphasize that the vector neural activity index $\text{NAI}(L)$, as defined in [210] and equation (3.45), is *different* from the power of the neural activity index beamformer $P_{\text{NAI}}(L)$ as it is defined in [135] and in definition 3.1.5 of this thesis. This (extremely confusing) naming convention originates from the fact that both [210] as well as [135] also define scalar versions of their localizer functions, i.e., localizer functions that apply to lead field vectors $l \in \mathbb{R}^{L \times 1}$. The scalar version of the neural activity index localizer now agrees for both authors. We thus see that [210] and [135] offer different contenders for the vector beamformer generalization of the neural activity index.

The usage of vector beamformers in the MEEG community is, in our opinion, currently lacking a set of clear recommendations. The main reason for this is that there are, to the best of our knowledge, at least 13 different candidates for vector beamformer localizer function construction strategies that are available in the literature. The standard reference book [190] does, as far as we can tell, not give any guidance on which one of the many vector beamformers to use in what situation. And while the recent review paper [231] at least gives a recommendation against some of these vector beamformers, their argument is based on these beamformers not being “rotationally invariant”. We think that while rotational invariance is certainly an interesting property, the central measure for the performance of a source reconstruction approach should be its ability to reconstruct sources. In the following, we also want to give a recommendation for vector beamformers. We will base this recommendation on whether the respective vector beamformer is biased, i.e., if it is able to reconstruct a single dipole. It turns out that out of the 13 vector beamformers we will now define, only 2 are unbiased.

We remind the reader that a beamformer is already determined from the mapping $L \mapsto M(L)$ of a lead field to its weighting matrix, as this by equation (3.13) determines the filter $W(L)$ and thus also all derived quantities. We will now investigate the following vector beamformer localizer functions.

1. The neural activity index and SAM beamformers with a power estimate of the first kind, as defined in [135] and definition 3.1.5.
2. The vector unit noise gain (UNG) beamformer from [190], as defined in [190], eq. (4.85). Given a lead field $L \in \mathbb{R}^{L \times 3}$, one first defines for $1 \leq i \leq 3$ the values

$$\gamma_i := e_i^T (L^T R^{-1} L)^{-1} \cdot L^T R^{-2} L \cdot (L^T R^{-1} L)^{-1} e_i \in \mathbb{R},$$

where e_i denotes the i -th unit vector, and then sets

$$M(L) := \begin{pmatrix} \frac{1}{\sqrt{\gamma_1}} & 0 & 0 \\ 0 & \frac{1}{\sqrt{\gamma_2}} & 0 \\ 0 & 0 & \frac{1}{\sqrt{\gamma_3}} \end{pmatrix}.$$

This is then proposed to be used with either a power estimate of the first or of the second kind.

3. The vector array gain (AG) beamformer from [190]. In fact, [190] defines two different vector AG beamformers, see equations (4.66) and (4.74) in that book. Let $L = (l_1, l_2, l_3) \in \mathbb{R}^{L \times 3}$ denote a lead field. The first proposed vector AG beamformer corresponds to the choice

$$M(L) := \begin{pmatrix} \|l_1\|_2 & 0 & 0 \\ 0 & \|l_2\|_2 & 0 \\ 0 & 0 & \|l_3\|_2 \end{pmatrix},$$

and the second one corresponds to the choice

$$M(L) := \|L\| \cdot I_3.$$

The authors in [190] do, as far as we can tell, not specify what matrix norm they use, and, in fact, there are different conventions in the MEEG community. The recent review paper [231] proposes to choose the Frobenius norm. Furthermore, [231] *wrongly* claims that FieldTrip also uses the Frobenius norm, when FieldTrip in fact uses the operator norm with respect to the euclidean norm, as can be seen by directly looking at the implementation.¹³ Thus, there are three candidates for the vector AG beamformer, each of which can be used with either the power estimate of the first or second kind, yielding a total of 6 localizer functions.

4. The unit gain beamformer (UG) as it is defined in [190], equation (4.58), and in definition 3.1.5, which can be used with either a power estimate of the first or second kind.
5. The vector neural activity index NAI(L) from [210] and equation (3.45).

The surprising result is now that out of these 13 localizers, only the neural activity index and SAM beamformers with the power estimate of the first kind, i.e., the localizer functions suggested in [135], are unbiased. The remaining 11 localizers, i.e., the vector NAI from [210] and all of the 10 vector beamforming strategies from [190] are biased, i.e., they are in general unable to correctly localize single dipoles. We will now prove this statement.

Theorem 3.3.5. *The neural activity index and SAM beamformers with the power estimate of the first kind, as they are defined in definition 3.1.5, are unbiased.*

Proof. If d_0 is the topography of the true dipole, the map $L \mapsto \text{GOF}(d_0, L)$ has a global maximum of 1 in the true lead field. By theorem 3.3.1, the power estimates of the first kind for the neural activity index and SAM beamformers are strictly monotonically increasing maps in this goodness of fit. Hence, they also attain a global maximum in the true lead field, which finishes the proof. \square

We will now show that the remaining 11 vector beamformers are biased. Because of its historical significance in the MEEG community, we will present an in-depth discussion of the vector neural activity index from [210], showing that its localizer function will almost never attain a global maximum in the true lead field. For the remaining 10 vector beamformers, we will contend ourselves with presenting examples where the global maximum condition is violated.

Theorem 3.3.6 (Bias of the original NAI). *The vector neural activity index from [210] is biased. Concretely, let the data d be described by a noisy single-source setting*

$$d = L_{x_0} \cdot M \cdot s + n,$$

where s is the source amplitude, M the source orientation, $L_{x_0} \in \mathbb{R}^{L \times 3}$ the lead field matrix, n the noise vector, L_{x_0} has full rank, and the source amplitude s is uncorrelated to the components of the noise vector n . Let $N^{-\frac{1}{2}} L_{x_0} = U S V^\top$ be the thin singular value decomposition, and let $S V^\top M = (v_1, v_2, v_3)$. Then, if the condition $|v_1| = |v_2| = |v_3|$ is false, we can find a lead field $\tilde{L} \in \mathbb{R}^{n \times 3}$ such that

$$\text{NAI}(\tilde{L}) > \text{NAI}(L_{x_0}),$$

where $\text{NAI}(L)$ is defined as in (3.45).

¹³see https://github.com/fieldtrip/fieldtrip/blob/master/inverse/ft_inverse_lcmv.m, line 370 at the point of writing.

Proof. Let $L \in \mathbb{R}^{L \times 3}$ denote an arbitrary lead field. In the proof of theorem 3.3.1 we computed

$$(L^\top R^{-1} L)^{-1} = (L^\top N^{-1} L)^{-1} + \frac{\mu_{d_0, N}}{\langle N^{-1} d_0, d_0 \rangle} \cdot \frac{(L^\top N^{-1} L)^{-1} L^\top N^{-1} d_0 (L^\top N^{-1} d_0)^\top (L^\top N^{-1} L)^{-1}}{1 - \mu_{d_0, N} \cdot \text{GOF}_{N^{-1}}(d_0, L)},$$

where $d_0 = \sqrt{\mathbb{E}[s^2]} \cdot L_0 \cdot \eta_0$. We thus have for an arbitrary lead field matrix $L \in \mathbb{R}^{n \times 3}$ that

$$\begin{aligned} \text{NAI}(L) &= \frac{\text{Tr} \left((L^\top R^{-1} L)^{-1} \right)}{\text{Tr} \left((L^\top N^{-1} L)^{-1} \right)} \\ &= 1 + \frac{\mu_{d_0, N}}{\langle N^{-1} d_0, d_0 \rangle} \cdot \frac{1}{1 - \mu_{d_0, N} \cdot \text{GOF}_{N^{-1}}(d_0, L)} \cdot \frac{(L^\top N^{-1} d_0)^\top (L^\top N^{-1} L)^{-2} L^\top N^{-1} d_0}{\text{Tr} \left((L^\top N^{-1} L)^{-1} \right)}. \end{aligned}$$

Since $N^{-\frac{1}{2}} L_{x_0} = U S V^\top$ is a thin singular value decompositions, a straightforward computation yields $(L_{x_0}^\top N^{-1} L_{x_0})^{-1} = V S^{-2} V^\top$, $(L_{x_0}^\top N^{-1} L_{x_0})^{-2} = V S^{-4} V^\top$, and thus

$$(N^{-\frac{1}{2}} L_{x_0}) (L_{x_0}^\top N^{-1} L_{x_0})^{-2} (N^{-\frac{1}{2}} L_{x_0})^\top = U S^{-2} U^\top$$

Using this, we get

$$\text{NAI}(L_{x_0}) = 1 + \frac{\mu_{d_0, N}}{\langle N^{-1} d_0, d_0 \rangle} \cdot \frac{1}{1 - \mu_{d_0, N} \cdot \text{GOF}_{N^{-1}}(d_0, L_{x_0})} \cdot \frac{(U^\top N^{-\frac{1}{2}} d_0)^\top S^{-2} (U^\top N^{-\frac{1}{2}} d_0)}{\text{Tr}(S^{-2})}.$$

From lemma 3.2.2, it is easy to see that $\text{GOF}_{N^{-1}}(d_0, L_{x_0})$ is independent of the diagonal entries of S . In particular, we can change the diagonal values of S without changing the goodness of fit.

Let $S = \text{diag}(s_1, s_2, s_3)$. Since L_{x_0} is by assumption full rank, we have $s_1, s_2, s_3 > 0$. Let $0 < \epsilon < \min\{s_1^{-2}, s_2^{-2}, s_3^{-2}\}$. Since, by assumption, the statement $|v_1| = |v_2| = |v_3|$ is false, we have indices i, j with $i \neq j$ and $|v_i| < |v_j|$. Denote by k the remaining index so that $\{i, j, k\} = \{1, 2, 3\}$. Now define $\tilde{s}_i = (s_i^{-2} - \epsilon)^{-\frac{1}{2}}$, $\tilde{s}_j = (s_j^{-2} + \epsilon)^{-\frac{1}{2}}$, $\tilde{s}_k = s_k$. Then let $\tilde{S} = \text{diag}(\tilde{s}_1, \tilde{s}_2, \tilde{s}_3)$. Then, we have $\text{Tr}(S^{-2}) = \text{Tr}(\tilde{S}^{-2})$, $\tilde{s}_i^{-2} = s_i^{-2} - \epsilon$, and $\tilde{s}_j^{-2} = s_j^{-2} + \epsilon$. Now notice that $U^\top N^{-\frac{1}{2}} d_0 = \sqrt{\mathbb{E}[s^2]} S V^\top M = \sqrt{\mathbb{E}[s^2]} \cdot (v_1, v_2, v_3)$. Using this, we get

$$\begin{aligned} (U^\top N^{-\frac{1}{2}} d_0)^\top S^{-2} (U^\top N^{-\frac{1}{2}} d_0) &= \mathbb{E}[s^2] \cdot \sum_{l=1}^3 s_l^{-2} v_l^2 \\ &< \mathbb{E}[s^2] \cdot \sum_{l=1}^3 \tilde{s}_l^{-2} v_l^2 \\ &= (U^\top N^{-\frac{1}{2}} d_0)^\top \tilde{S}^{-2} (U^\top N^{-\frac{1}{2}} d_0). \end{aligned}$$

If we now define $\tilde{L} = N^{\frac{1}{2}} U \tilde{S} V^\top$, we get

$$\begin{aligned} \text{NAI}(L_{x_0}) &= 1 + \frac{\mu_{d_0, N}}{\langle N^{-1} d_0, d_0 \rangle} \cdot \frac{1}{1 - \mu_{d_0, N} \cdot \text{GOF}_{N^{-1}}(d_0, L_{x_0})} \cdot \frac{(U^\top N^{-\frac{1}{2}} d_0)^\top S^{-2} (U^\top N^{-\frac{1}{2}} d_0)}{\text{Tr}(S^{-2})} \\ &< 1 + \frac{\mu_{d_0, N}}{\langle N^{-1} d_0, d_0 \rangle} \cdot \frac{1}{1 - \mu_{d_0, N} \cdot \text{GOF}_{N^{-1}}(d_0, \tilde{L})} \cdot \frac{(U^\top N^{-\frac{1}{2}} d_0)^\top \tilde{S}^{-2} (U^\top N^{-\frac{1}{2}} d_0)}{\text{Tr}(\tilde{S}^{-2})} \\ &= \text{NAI}(\tilde{L}). \end{aligned}$$

Thus, we have found a lead field matrix \tilde{L} such that $\text{NAI}(L_{x_0}) < \text{NAI}(\tilde{L})$. In fact, since the map $L \mapsto \text{NAI}(L)$ is continuous, we can find a small perturbation L of this \tilde{L} such that still $\text{NAI}(L_{x_0}) < \text{NAI}(L)$ and, additionally, $\text{GOF}_{N-1}(d_0, L) < \text{GOF}_{N-1}(d_0, L_{x_0})$. Hence, we can even find a lead field matrix that simultaneously has a higher value for NAI and a worse fit to the data than L_{x_0} . \square

Theorem 3.3.7 (Bias of the vector beamformers from [190]). *The vector UG, UNG, and AG beamformers are biased, both when using the power estimate of the first kind as well as when using the power estimate of the second kind.*

Proof. In [190], the author always assumes white noise. Hence, we will show that vector UG, UNG, and AG beamformers are biased even under the assumption of white noise. In the following, assume that the noise covariance matrix is given by $N = \text{Id}_L$. Furthermore, we assume that for the expected source power we have $Q = \sqrt{\mathbb{E}[s^2]} = 1$.

We start with the vector UNG beamformer. Then define

$$L_{x_0}^{\text{UNG}} = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 0 \end{pmatrix}, \tilde{L}^{\text{UNG}} = \begin{pmatrix} 1 & 0.5 & 0.5 \\ 0 & 1 & -0.5 \\ 0 & 0 & 1 \\ 0.1 & 0 & 0 \\ 0 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 0 \end{pmatrix},$$

and $M = (1, 0, 0)$. Then the UNG power of the first kind for the true lead field $L_{x_0}^{\text{UNG}}$ is 3.5, and the UNG power of the second kind is 2.0. The UNG power of the first kind for the lead field \tilde{L}^{UNG} , on the other hand, is ca. 3.53, and the UNG power of the second kind is ca. 2.28. Hence, the vector UNG beamformer is biased for both kinds of power estimates.

We continue with the vector AG beamformer for the choice

$$M(L) := \begin{pmatrix} \|l_1\|_2 & 0 & 0 \\ 0 & \|l_2\|_2 & 0 \\ 0 & 0 & \|l_3\|_2 \end{pmatrix},$$

In this case, one can, e.g., choose

$$L_{x_0,1}^{\text{AG}} = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 0 \end{pmatrix}, \tilde{L}_1^{\text{AG}} = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0.2 & 0 & 0 \\ 0 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 0 \end{pmatrix},$$

and $M = (1, 0, 0)$. Then, the true power of the first kind is 6 and the true power of the second kind is 4. For \tilde{L}_1^{AG} , on the other hand, the power of the first kind is ca. 10.39 and the power of the second kind is ca. 8.39.

For the vector AG beamformer with the choice $M(L) := \|L\|_{\text{frobenius}} \cdot I_3$ and with the choice $M(L) := \|L\|_2 \cdot I_3$, where $\|L\|_2 := \sup_{\|x\|_2=1} \|Lx\|_2$, one can in both cases choose

$$L_{x_0,2}^{\text{AG}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 0 \end{pmatrix}, \tilde{L}_2^{\text{AG}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0.2 & 0 & 0 \\ 0 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 0 \end{pmatrix},$$

and $M = (1, 0, 0)$. In the Frobenius case, the true power of the first kind is 12 and the true power of the second kind is 6. For \tilde{L}_2^{AG} , the power of the first kind is ca. 18.74 and the power of the second kind is ca. 11.19. In the operator norm case, the true power of the first kind is 4, and the power of the second kind is 2. For \tilde{L}_2^{AG} , the power of the first kind is ca. 12.41, and the power of the second kind is ca. 7.41.

Finally, for the vector UG beamformer, one can choose

$$L_{x_0}^{\text{UG}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 0 \end{pmatrix}, \tilde{L}^{\text{UG}} = \begin{pmatrix} 0.1 & 0 & 0 \\ 0 & 0.1 & 0 \\ 0 & 0 & 0.1 \\ 0.5 & 0 & 0 \\ 0 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 0 \end{pmatrix},$$

and $M = (1, 0, 0)$. Then, the true power of the first kind is 4 and the true power of the second kind is 2. For \tilde{L}^{UG} , the power of the first kind is ca. 203.92, and the power of the second kind is 100. Note that this last example illustrates the aforementioned “center of head bias” of the UG beamformer, where small entries of a lead field lead to a large output filter power. \square

Given how easy it is to construct examples where the vector beamformers from [190] fail to correctly reconstruct single dipoles, we find it surprising that these beamformers are still used in the MEEG community, especially as there are other beamformers, e.g., the ones defined in [135] or the scalar beamformers from [190], that do not have this shortcoming.

This finishes our discussion on the bias of vector beamformers.

3.3.2 Reconstructing source orientations with beamforming

In this chapter, we will focus on so-called *scalar beamformers*. In the MEEG community, one speaks of a scalar beamformer if the source configuration used in the computation of the beamformer filter in equation (3.13) consists of a single dipole, i.e., $L \in \mathbb{R}^{L \times 1}$, and then also $W(L) \in \mathbb{R}^{L \times 1}$. In accordance with the literature, and to differentiate the scalar from the non-scalar case, we will in this subsection always use lowercase letters for lead fields and filters, e.g., $l \in \mathbb{R}^{L \times 1}$ for a generic lead field and $w \in \mathbb{R}^{L \times 1}$ for a generic filter.

The focus in this section will be on investigating algorithms for the reconstruction of source orientations using scalar beamformers. As a first step, we will introduce a general strategy for

estimating orientations using beamforming, based on [190]. The idea is, in principle, the same as the one underlying the power estimate of the second kind in the preceding subsection.

Definition 3.3.8 (General scalar beamformer orientation reconstruction strategy). Let d denote the measurement vector, which we again assume to be a random variable. Let x_0 be an arbitrary source position. For $M \in \mathbb{R}^3$ denote by $l_{x_0, M} \in \mathbb{R}^L$ the lead field generated by the dipole $M \cdot \delta_{x_0}$. Let $w(l_{x_0, M}) \in \mathbb{R}^L$ denote the beamformer corresponding to the weighting $m(l_{x_0, M}) \in \mathbb{R}$. Then, we estimate the orientation $M_{\text{opt}} \in \mathbb{R}^3$ of a source at position x_0 via

$$M_{\text{opt}} = \arg \max_{\|M\|=1} \mathbb{E} [\langle w(l_{x_0, M}), d \rangle^2].$$

If again $R = \mathbb{E}[d \cdot d^\top]$ denotes the covariance matrix, a straightforward computation using lemma 3.15 implies

$$M_{\text{opt}} = \arg \max_{\|M\|=1} \frac{m(l_{x_0, M})^2}{\langle l_{x_0, M}, R^{-1} l_{x_0, M} \rangle}. \quad (3.46)$$

In the following, we will always assume a fixed position x_0 , and will hence suppress it in the notation, i.e., we simply write l_M instead of $l_{x_0, M}$. Contrary to the vector beamformer case, there are only three common choices for scalar beamformers. Concretely, these are the UG, the NAI, and the SAM beamformers, as they are defined in definition 3.1.5.¹⁴ An interesting observation, that is typically attributed to [191], is that the optimization problem (3.46) can be reduced to an eigenvalue problem.

Lemma 3.3.9. *Let $A, B \in \mathbb{R}^{k \times k}$ be symmetric positive definite matrices. Then $B^{-1} \cdot A$ is diagonalizable with real eigenvalues and eigenvectors, and if λ_{\max} denotes the maximal eigenvalue of $B^{-1} \cdot A$ and v_{\max} is an arbitrary normed eigenvector to this eigenvalue, we have*

$$v_{\max} = \arg \max_{\|v\|=1} \frac{\langle v, Av \rangle}{\langle v, Bv \rangle},$$

and the maximum is only attained in the eigenvectors to the eigenvalue λ_{\max} .

Proof. Note that $\langle x, y \rangle_B := \langle Bx, y \rangle$ defines an inner product on \mathbb{R}^k , and that the linear map given by $x \mapsto B^{-1} \cdot A \cdot x$ defines a self-adjoint and positive definite operator on the finite-dimensional Hilbert space $(\mathbb{R}^k, \langle \cdot, \cdot \rangle_B)$. By the spectral theorem, there is thus an eigenbasis of \mathbb{R}^k of eigenvectors of $B^{-1} \cdot A$, that is orthogonal with respect to $\langle \cdot, \cdot \rangle_B$. Let v_1, \dots, v_k denote such an eigenbasis, and let $\lambda_1, \dots, \lambda_k$ denote the corresponding positive eigenvalues. Then, an arbitrary $v \in \mathbb{R}^k$ can be decomposed in this basis as $v = \sum_{i=1}^k \mu_i v_i$, and we have

$$\frac{\langle v, Av \rangle}{\langle v, Bv \rangle} = \frac{\sum_{i=1}^k \lambda_i \cdot \mu_i^2}{\sum_{i=1}^k \mu_i^2} \leq \lambda_{\max}.$$

Thus λ_{\max} is an upper bound for $\frac{\langle v, Av \rangle}{\langle v, Bv \rangle}$. If now v_{\max} is an arbitrary eigenvector to this eigenvalue, it is easy to see that the value λ_{\max} is attained in this eigenvector. Furthermore, the computation above shows that the inequality is strict if v has a non-zero component in a direction whose eigenvalue is smaller than λ_{\max} . \square

¹⁴Remember that this includes the scalar AG and UNG beamformers as special cases for white noise.

Note that instead of looking for eigenvectors of $B^{-1} \cdot A$, one could equivalently search for vectors v fulfilling $A \cdot v = \lambda \cdot B \cdot v$, which is typically called a *generalized eigenvalue problem* for the pair (A, B) . In this case, v is called a *generalized eigenvector*, and λ is a generalized eigenvalue of the pair (A, B) . Additionally, note that if v is a generalized eigenvector with generalized eigenvalue λ for the pair (A, B) , then v is also a generalized eigenvector with the generalized eigenvalue $\frac{1}{\lambda}$ for the pair (B, A) . Let $\lambda_{\min}(B, A)$ denote the minimum generalized eigenvalue of the pair (B, A) , and let $\nu_{\min}(B, A)$ denote the set of all generalized eigenvectors with respect to this generalized eigenvalue. Then, lemma 3.3.9 implies that

$$\nu_{\min}(B, A) = \arg \max_{v \neq 0} \frac{\langle v, Av \rangle}{\langle v, Bv \rangle}. \quad (3.47)$$

Equation (3.47) is the formulation of lemma 3.3.9 that is typically used in MEEG beamforming. Furthermore, we define $\nu_{\min}^n(B, A) := \{v \in \nu_{\min}(B, A) \mid \|v\|_2 = 1\}$ to be the set of normed generalized eigenvectors of the pair (B, A) corresponding to the the smallest generalized eigenvalue. We can then state equation (3.47) equivalently as

$$\nu_{\min}^n(B, A) = \arg \max_{\|v\|_2=1} \frac{\langle v, Av \rangle}{\langle v, Bv \rangle}. \quad (3.48)$$

Remark 3.3.10. We now want to explicitly express the orientation estimate of the scalar UG, NAI, and SAM beamformers as solutions of generalized eigenvalue problems. Let $L \in \mathbb{R}^{L \times 3}$ denote the lead field at position x_0 . As shown before, we then have $l_M = L \cdot M \in \mathbb{R}^L$ for each $M \in \mathbb{R}^3$. We refer to definition 3.1.5 for the definitions of UG, NAI, and SAM beamformers.

1. The scalar UG beamformer corresponds to the choice $m(l_M) = 1$, which by equation (3.46) yields for the optimum orientation $M_{\text{UG}} := M_{\text{opt}}$ that

$$M_{\text{UG}} = \arg \max_{\|M\|=1} \frac{1}{\langle M, (L^\top R^{-1} L) \cdot M \rangle} = \nu_{\min}^n(L^\top R^{-1} L, \text{Id}_3). \quad (3.49)$$

Note that a generalized eigenvector of the pair $(L^\top R^{-1} L, \text{Id}_3)$ is the same as an eigenvector of $L^\top R^{-1} L$. The UG orientation estimate is thus given by a normalized eigenvector of $L^\top R^{-1} L$ with respect to the smallest eigenvalue.

2. The scalar NAI beamformer corresponds to the choice $m(l_M) = \sqrt{\langle N^{-1} l_M, l_M \rangle}$. By equation (3.46), this yields for the optimum orientation $M_{\text{NAI}} := M_{\text{opt}}$ that

$$M_{\text{NAI}} = \arg \max_{\|M\|=1} \frac{\langle M, (L^\top N^{-1} L) \cdot M \rangle}{\langle M, (L^\top R^{-1} L) \cdot M \rangle} = \nu_{\min}^n(L^\top R^{-1} L, L^\top N^{-1} L). \quad (3.50)$$

In the special case of white noise, i.e., for the array gain beamformer, we get the orientation estimate

$$M_{\text{AG}} = \nu_{\min}^n(L^\top R^{-1} L, L^\top L). \quad (3.51)$$

3. The scalar SAM beamformer corresponds to the choice $m(l_M) = \frac{\langle l_M, R^{-1} \cdot l_M \rangle}{\sqrt{\langle l_M, R^{-1} N R^{-1} l_M \rangle}}$. By equation (3.46), this yields for the optimum orientation $M_{\text{SAM}} := M_{\text{opt}}$ that

$$M_{\text{SAM}} = \arg \max_{\|M\|=1} \frac{\langle M, (L^\top R^{-1} L) \cdot M \rangle}{\langle M, (L^\top R^{-1} N R^{-1} L) \cdot M \rangle} = \nu_{\min}^n(L^\top R^{-1} N R^{-1} L, L^\top R^{-1} L). \quad (3.52)$$

In the special case of white noise, i.e., for the unit noise gain beamformer, we get the orientation estimate

$$M_{\text{UNG}} = \nu_{\min}^n (L^\top R^{-2} L, L^\top R^{-1} L). \quad (3.53)$$

In the way we have presented beamformers, it is obvious that the UNG beamformer is a special case of the SAM beamformer, and that the AG beamformer is a special case of the NAI beamformer. If one, however, looks at the recent MEEG beamforming literature, one sees that many researchers in the MEEG community are not aware of this relation. In all recent publications that we are aware of, such as e.g., [231], NAI and SAM beamformers are treated as explicitly different from AG and UNG beamformers. Furthermore, while the eigendecomposition characterizations for the AG and UNG beamformer orientations in (3.51) and (3.53) are well known in the community, probably because they are explicitly derived and discussed in the standard reference work [190], the corresponding characterizations for the NAI and SAM beamformers in (3.50) and (3.52) are not. In fact, in personal communication, I have multiple times encountered the belief that while AG and UNG beamformer orientations can be estimated from an eigendecomposition, the SAM orientation cannot and needs to be estimated from a non-linear optimizer. As a concrete example, the FieldTrip implementation of the SAM beamformer did not use an eigendecomposition, but instead a call to an optimization toolbox applied on the SAM power functional to compute an orientation estimate. During the writing process of [30], we realized the connection between SAM and UNG beamformers and reimplemented the orientation estimation code in FieldTrip. This yielded a speedup of a factor of $\times 1.7$ for FieldTrips' SAM beamformer implementation, and even slightly increased the accuracy of orientation estimates, to within the tolerance of the previously used optimizer.¹⁵ We were, however, not the first to notice (3.52), as for example this¹⁶ presentation by Stephen Robinson demonstrates.

If we want to reconstruct orientations using scalar beamformers, the first question is, of course, just as in the preceding subsection, how well this works. In [30], we approached this question by investigating a noisy single-source setting, i.e., we modelled the data d as a random vector given by

$$d = L_{x_0} \cdot M_0 \cdot s + n,$$

with the same assumptions as, e.g., in theorem 3.3.6, and tried to reconstruct the orientation M_0 . We, in particular, assumed that we already know the true source location x_0 . Before continuing, note that the orientation estimates, as they are defined in remark 3.3.10, are, at best, only determined up to sign, i.e., if v is a beamforming orientation estimate, then so is $-v$. This is analogous to dipole scans, where one has $\text{GOF}_C(d, l) = \text{GOF}_C(d, -l)$ for each lead field vector l . Furthermore, let $\mathbb{R}^* := \mathbb{R} \setminus \{0\}$ denote the non-zero real numbers.

Definition 3.3.11. We call an orientation estimation strategy *unbiased* if, given a noisy single-source setting

$$d = L_{x_0} \cdot M_0 \cdot s + n$$

with $M_0 \neq 0$, it produces an orientation estimate contained in $\mathbb{R}^* \cdot M_0$.

Theorem 3.3.12. *Estimating the orientation using either a NAI beamformer, as defined in (3.50), or a SAM beamformer, as defined in (3.52), is unbiased.*

¹⁵See <https://github.com/fieldtrip/fieldtrip/pull/2325> for the corresponding pull request and https://github.com/fieldtrip/fieldtrip/blob/master/inverse/ft_inverse_sam.m for the current implementation.

¹⁶https://megcore.nih.gov/images/9/9c/StephenRobinson_SAM_Beamforming_Talk_10222021.pdf, slide 12

Proof. Let M_0 denote the true source orientation. By theorem 3.3.1, the NAI and SAM orientation estimates are given by normalized directions maximizing the map

$$M \mapsto \text{GOF}_{N^{-1}}(L_{x_0} \cdot M_0, L_{x_0} \cdot M).$$

The maximizers of this map are obviously $M = \pm \frac{M_0}{\|M_0\|_2}$. \square

For comparison, we derived the same theorem in [30], but as we did not know of theorem 3.3.1 at the time of publishing this paper, we proved the theorem by explicitly computing the corresponding generalized eigendecompositions. We find it remarkable how much simpler the proof is rendered by the identification of NAI and SAM beamformers with dipole scanning. Furthermore, this identification also gives a theoretical reasoning for why we were unable to work out a benefit of either SAM over NAI beamforming or of NAI over SAM beamforming, as in a noisy single-source setting, these two beamformers are just bijective transformations of one another.

Remark 3.3.13. One central aspect, which is implicitly contained in theorem 3.3.12, is that an unbiased orientation reconstruction requires the knowledge of the noise covariance matrix N . Remember that N enters the definition of the NAI and SAM beamformers via the definition of the weights, see, e.g., remark 3.3.10. To illustrate this point, we will present an example where, in the presence of non-white noise, the AG beamformer and the UNG beamformer both fail to correctly reconstruct the orientation. Assume that the first sensor is noisier than the other sensors, with a noise covariance matrix given by $N = \text{diag}(3, 1, \dots, 1) \in \mathbb{R}^{L \times L}$. Furthermore, assume that the true lead field is given by

$$L = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \\ \vdots & \vdots & \vdots \\ 0 & 0 & 0 \end{pmatrix} \in \mathbb{R}^{L \times 3},$$

and that the true orientation is given by $e_2 = (0, 1, 0)$. If we furthermore assume $\mathbb{E}[s^2] = 1$, we have

$$(L^\top L)^{-1} \cdot (L^\top R^{-1} L) = (L^\top R^{-1} L)^{-1} \cdot (L^\top R^{-2} L) = \begin{pmatrix} \frac{1}{3} & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Hence, both the AG as well as the UNG beamformer reconstruct $e_1 = (1, 0, 0)$, while the true orientation is $e_2 = (0, 1, 0)$.

We now continue with the aforementioned combination of EEG and MEG. In [30], we proposed to do this as follows. First, we create an initial estimate of the source orientation from the EEG data. Then, we extract the “tangential” component of this reconstruction and replace it with an estimation based on MEG data. The heuristic behind this approach is that for a given MEG lead field $L_{x_0} \in \mathbb{R}^{L \times 3}$ there is typically a direction $\eta \in \mathbb{R}^3$ with the property that $\|L_{x_0} \cdot \eta\|$ is much smaller than $\|L_{x_0} \cdot \xi\|$ for all $\xi \in \{\eta\}^\perp$ with $\|\eta\| \approx \|\xi\|$. Furthermore, this direction η is typically oriented roughly normal to the skull surface, and is thus called *radial* or *quasi-radial* in the MEEG community [56].¹⁷ It follows that the MEG is less sensitive to the orientation

¹⁷The existence of such an η is commonly explained by referring to [178], where it is shown that in a multilayer-sphere model the B -field generated by a radial dipole vanishes outside of the model. It is then argued that the human head looks roughly like a multilayer sphere.

component in $\mathbb{R} \cdot \eta$ than to the orientation component in $\{\eta\}^\perp$. It is thus a natural idea to first estimate the complete orientation, including the component in the direction of η , by using the EEG, and then refine the estimate in the orthogonal plane $\{\eta\}^\perp$ by using the MEG.

In [30], we were able to show that, under the assumption of white noise, the AG and UNG beamformers are able to reconstruct the tangential component of the true source orientation by only searching for orientations in the tangential plane. At the time of writing [30], we did not realize how sensitive this statement is to the assumption of white noise. On deeper investigation, it turns out that the meaning of what the “tangential” plane even is has to be chosen in accordance with the noise covariance matrix N . The common strategy of just estimating the “tangential” plane from a singular value decomposition of the lead field L_{x_0} is, in fact, generally insufficient. We will now present a generalization of our result from [30], which is applicable for general noise covariance matrices. Furthermore, we will present an example where, in the presence of non-white noise, the estimation of the tangential orientation component using NAI and SAM beamformers fails for the naive notion of the “tangential” plane.

Lemma 3.3.14. *Let $A \in \mathbb{R}^{L \times k}$ be full rank with $k < L$. Assume a noisy single-source setting, i.e.*

$$d = L_{x_0} \cdot M_0 \cdot s + n,$$

and assume that $L_{x_0} \cdot M_0$ is not N^{-1} -orthogonal to $\text{Im}(A)$. Then

$$\begin{aligned} \nu_{\min}(A^\top R^{-1} A, A^\top N^{-1} A) &= \nu_{\min}(A^\top R^{-1} N R^{-1} A, A^\top R^{-1} A) \\ &= \mathbb{R}^* \cdot (A^\top N^{-1} A)^{-1} A^\top N^{-1} L_{x_0} \cdot M_0. \end{aligned}$$

Proof. By theorem 3.3.1 and 3.3.9, we have

$$\nu_{\min}(A^\top R^{-1} A, A^\top N^{-1} A) = \nu_{\min}(A^\top R^{-1} N R^{-1} A, A^\top R^{-1} A) = \arg \max_{0 \neq \eta \in \mathbb{R}^k} \text{GOF}_{N^{-1}}(L_{x_0} \cdot M_0, A \cdot \eta).$$

By remembering that the N^{-1} -orthogonal projection onto $\text{Im}(A)$ is given by

$$p(v) = A \cdot (A^\top N^{-1} A)^{-1} \cdot A^\top N^{-1} \cdot v,$$

the result follows. \square

Note that by choosing $A = L_{x_0}$, it follows that lemma 3.3.14 contains theorem 3.3.12 as a special case.

Now, again assume we have a noisy single-source setting with true lead field $L_{x_0} \in \mathbb{R}^{L \times 3}$ and true orientation $M_0 \in \mathbb{R}^3$, i.e., the data is given by $d = L_{x_0} \cdot M_0 \cdot s + n$. Let $W \subset \mathbb{R}^3$ denote a 2-dimensional subspace. Our goal is now to investigate under what conditions it is possible to restrict the NAI and SAM orientation reconstruction procedures to the space W , i.e., under what conditions the expected beamformer power output over the space W is maximized in the linear space spanned by the orthogonal projection of the true orientation onto the space W .

If $\{w_1, w_2\}$ denotes an orthonormal basis of W , $\tilde{V} := (w_1, w_2)$, and $L_{\tilde{V}} := L_{x_0} \cdot \tilde{V} \in \mathbb{R}^{L \times 2}$, it follows from lemma 3.3.9 that the restriction to the subspace W is possible if, and only if, one has

$$\nu_{\min}(L_{\tilde{V}}^\top R^{-1} L_{\tilde{V}}, L_{\tilde{V}}^\top N^{-1} L_{\tilde{V}}) = \nu_{\min}(L_{\tilde{V}}^\top R^{-1} N R^{-1} L_{\tilde{V}}, L_{\tilde{V}}^\top R^{-1} L_{\tilde{V}}) = \mathbb{R}^* \cdot \begin{pmatrix} \langle w_1, M_0 \rangle \\ \langle w_2, M_0 \rangle \end{pmatrix}.$$

By lemma 3.3.14, this is in turn equivalent to

$$\mathbb{R}^* \cdot \left(L_{\tilde{V}}^\top N^{-1} L_{\tilde{V}} \right)^{-1} L_{\tilde{V}}^\top N^{-1} L_{x_0} \cdot M_0 = \mathbb{R}^* \cdot \begin{pmatrix} \langle w_1, M_0 \rangle \\ \langle w_2, M_0 \rangle \end{pmatrix}.$$

Theorem 3.3.15. *Assume a noisy single-source setting for the MEG lead field $L_{x_0} \in \mathbb{R}^{L \times 3}$ and true orientation $M_0 \in \mathbb{R}^3$. Furthermore, let*

$$N^{-\frac{1}{2}} \cdot L_{x_0} = U S V^\top$$

with $U \in \mathbb{R}^{L \times 3}$, $S \in \mathbb{R}^{3 \times 3}$, and $V \in \mathbb{R}^{3 \times 3}$ denote a thin singular value decomposition. Let $S = \text{diag}(s_1, s_2, s_3)$, and assume $s_3 < s_2 \leq s_1$. Let $V = (v_1, v_2, v_3)$. We then define the tangential plane via

$$T_{L_{x_0}, N} := \text{span}\{v_1, v_2\}.$$

We assume that M_0 is not orthogonal to the tangential plane $T_{L_{x_0}, N}$, i.e., that there exists a tangential component that we can reconstruct. If then $p_T : \mathbb{R}^3 \rightarrow T_{L_{x_0}, N}$ denotes the orthogonal projection with respect to the standard inner product, we have

$$\arg \max_{0 \neq \eta \in T_{L_{x_0}, N}} P_{NAI}(l_\eta) = \arg \max_{0 \neq \eta \in T_{L_{x_0}, N}} P_{SAM}(l_\eta) = \mathbb{R}^* \cdot p_T(M_0).$$

Plainly speaking, we can thus reconstruct the tangential component of the true orientation by maximizing the projected beamformer power over the tangential plane.

Proof. Let $\tilde{V} = (v_1, v_2)$. By the discussion preceding the statement of the theorem, it suffices to show that

$$\mathbb{R}^* \cdot \left(L_{\tilde{V}}^\top N^{-1} L_{\tilde{V}} \right)^{-1} L_{\tilde{V}}^\top N^{-1} L_{x_0} \cdot M_0 = \mathbb{R}^* \cdot \begin{pmatrix} \langle v_1, M_0 \rangle \\ \langle v_2, M_0 \rangle \end{pmatrix},$$

where $L_{\tilde{V}} = L \cdot (v_1, v_2)$. Now, note that

$$L_{\tilde{V}}^\top N^{-1} L_{\tilde{V}} = \begin{pmatrix} s_1 & 0 \\ 0 & s_2 \end{pmatrix},$$

and

$$L_{\tilde{V}}^\top N^{-1} L_{x_0} \cdot M_0 = \begin{pmatrix} s_1 & 0 \\ 0 & s_2 \end{pmatrix} \cdot \begin{pmatrix} \langle v_1, M_0 \rangle \\ \langle v_2, M_0 \rangle \end{pmatrix}.$$

This finishes the proof. \square

The reader is invited to compare the proof of theorem 3.3.15 to the proof of theorem 1 in [30]. We find it again astonishing how much simpler, and also conceptually clearer, the argument is rendered by theorem 3.3.1.

Remark 3.3.16. For the validity of theorem 3.3.15, it is crucial that the tangential plane is derived from the singular value decomposition of the *pre-whitened* lead field $N^{-\frac{1}{2}} \cdot L_{x_0}$. If we were to instead compute the singular value decomposition of L_{x_0} , say

$$L_{x_0} = \hat{U} \cdot \hat{S} \cdot \hat{V}^\top,$$

with $\hat{U} \in \mathbb{R}^{L \times 3}$, $\hat{S} \in \mathbb{R}^{3 \times 3}$, and $\hat{V} = (\hat{v}_1, \hat{v}_2, \hat{v}_3) \in \mathbb{R}^{3 \times 3}$, and define a “tangential” plane via $\hat{T} = \text{span}(\hat{v}_1, \hat{v}_2) \subset \mathbb{R}^3$, the conclusion of the theorem will in general not hold. We want to present an example of this. Let

$$L_{x_0} := \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \\ \vdots & \vdots & \vdots \\ 0 & 0 & 0 \end{pmatrix} \in \mathbb{R}^{L \times 3}, N = \begin{pmatrix} 1 & 0 & 0 & & \\ 0 & 2 & 1 & & 0_{3,L-3} \\ 0 & 1 & 2 & & \\ & & & 0_{L-3,3} & \text{Id}_{L-3,L-3} \end{pmatrix} \in \mathbb{R}^{L \times L}, M_0 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \in \mathbb{R}^3,$$

and $\mathbb{E}[s^2] = 1$. Then the singular value decomposition of L_{x_0} is given by $\hat{U} = (e_1, e_2, e_3)$, $\hat{S} = \text{diag}(2, 2, 1)$, and $\hat{V} = (e_1, e_2, e_3)$. The tangential plane derived from this singular value decomposition is then given by $\mathbb{R}^2 \times \{0\}$. The direction defined by the orthogonal projection of M_0 on this tangential plane is thus given by $\mathbb{R} \cdot (1 \ 1 \ 0)^\top$. The NAI and SAM powers on this plane are, however, maximized in the direction $\mathbb{R} \cdot (\frac{4}{5} \ \frac{3}{5} \ 0)^\top$. Thus, maximizing the NAI and SAM powers on the “tangential” plane $\mathbb{R}^2 \times \{0\}$ fails to reconstruct the tangential component of the true orientation.

Before finishing up this section, we want to present one final result. We have now seen that NAI and SAM beamformers enable an unbiased estimation of the orientation of a neural source. We have, however, up until now, not talked about the UG beamformer. In equation (3.49), we derived an eigendecomposition-based characterization for the UG orientation estimate. It turns out that estimating source orientations using the scalar UG beamformers is biased. One can, however, show that, as long as the SNR in the measurement is sufficiently high, one can still expect the UG orientation estimate to be accurate. If one, however, applies UG beamforming to a low SNR scenario, the UG orientation estimate is strongly distorted. We presented a corresponding proof in [30], which we also want to discuss here.

Before stating the theorem, we want to remind the reader that the UG orientation estimate from equation (3.49) is normalized and only determined up to sign, i.e., if M_{UG} is an orientation estimate, then so is $-M_{UG}$. When we now write $M_{UG}(t) \xrightarrow{t \rightarrow t_0} x$, we mean that as soon as t is sufficiently close to t_0 , one has that either $M_{UG}(t)$ or $-M_{UG}(t)$ is close to $\frac{x}{\|x\|_2}$. Equivalently, if θ_t denotes the angle between $M_{UG}(t)$ and x , then $M_{UG}(t) \rightarrow x$ if, and only if, $\sin(\theta_t) \rightarrow 0$.

Theorem 3.3.17. *We assume a noisy single-source scenario, i.e., we model the data vector d as*

$$d = L_{x_0} \cdot M_0 \cdot s + n,$$

where $L_{x_0} \in \mathbb{R}^{L \times 3}$ has full rank, $M_0 \neq 0$, $Q^2 := \mathbb{E}[s^2] > 0$, s is uncorrelated to the components of n , n is zero mean, and the noise covariance matrix $N := \mathbb{E}[n \cdot n^\top] \in \mathbb{R}^{L \times L}$ is invertible. Let M_{UG} denote the UG orientation estimate, i.e., if $R := \mathbb{E}[d \cdot d^\top]$ is the signal covariance matrix, we have by equation (3.49) that

$$M_{UG} = \nu_{\min}(L_{x_0}^\top R^{-1} L_{x_0}).$$

We want to investigate how M_{UG} behaves with respect to N . Concretely, we keep all other parameters fixed and want to investigate how the UG orientation estimate behaves as a function of the noise covariance matrix N . We then have the following.

1. As $N \rightarrow 0$, we have that $M_{UG} \rightarrow M_0$.

2. Let $N_1 \in \mathbb{R}^{L \times L}$ be an arbitrary positive definite matrix. We assume that the smallest eigenvalue of $L_{x_0}^\top N_1^{-1} L_{x_0}$ is simple, and denote by v_{\min} a corresponding eigenvector. For $\rho > 0$ we define $N_\rho := \rho \cdot N_1$, and denote by $M_{UG,\rho}$ the UG orientation estimate corresponding to the noise covariance matrix N_ρ . Then, as $\rho \rightarrow \infty$, we have $M_{UG,\rho} \rightarrow v_{\min}$.

As a special case of 2), we get that, in a white noise scenario, the UG orientation estimate converges, for low SNR, towards the right singular vector of the lead field L_{x_0} corresponding to the smallest singular value.

Proof. At the beginning of the proof of theorem 3.3.1, we derived the equation

$$L_{x_0}^\top R^{-1} L_{x_0} = L_{x_0}^\top N^{-1} L_{x_0} - \frac{\mu_{d_0, N}}{\langle N^{-1} d_0, d_0 \rangle} (L_{x_0}^\top N^{-1} d_0) (L_{x_0}^\top N^{-1} d_0)^\top,$$

where $\mu_{d_0, N} = \frac{\langle N^{-1} d_0, d_0 \rangle}{1 + \langle N^{-1} d_0, d_0 \rangle}$. Now, by definition we have $d_0 = Q L_{x_0} M_0$, and if we thus define $\eta = Q \cdot M_0 \in \mathbb{R}^3$, this last equation can equivalently be written as

$$L_{x_0}^\top R^{-1} L_{x_0} = L_{x_0}^\top N^{-1} L_{x_0} - \frac{(L_{x_0}^\top N^{-1} L_{x_0}) \eta \cdot \eta^\top (L_{x_0}^\top N^{-1} L_{x_0})}{1 + \eta^\top (L_{x_0}^\top N^{-1} L_{x_0}) \eta}.$$

Now, note that the right-hand side of this equation has exactly the structure of a right-hand side in the Sherman-Morrison formula, which yields

$$L_{x_0}^\top R^{-1} L_{x_0} = \left((L_{x_0}^\top N^{-1} L_{x_0})^{-1} + \eta \cdot \eta^\top \right)^{-1},$$

which implies

$$M_{UG} = \nu_{\min}(L_{x_0}^\top R^{-1} L_{x_0}) = \nu_{\max}((L_{x_0}^\top N^{-1} L_{x_0})^{-1} + \eta \cdot \eta^\top).$$

We will first investigate the case when $N \rightarrow 0$. As a first step, we want to show that then also $(L_{x_0}^\top N^{-1} L_{x_0})^{-1} \rightarrow 0$. As $(L_{x_0}^\top N^{-1} L_{x_0})^{-1}$ is positive definite, it suffices to show that the largest eigenvalue of this matrix approaches zero. The largest eigenvalue of $(L_{x_0}^\top N^{-1} L_{x_0})^{-1}$ is the reciprocal of the smallest eigenvalue of $L_{x_0}^\top N^{-1} L_{x_0}$. By the same argument as in lemma 3.3.9, we now have

$$\begin{aligned} \lambda_{\min}((L_{x_0}^\top N^{-1} L_{x_0})) &= \min_{0 \neq \xi} \frac{\langle \xi, (L_{x_0}^\top N^{-1} L_{x_0}) \cdot \xi \rangle}{\langle \xi, \xi \rangle} = \min_{0 \neq \xi} \frac{\langle L_{x_0} \xi, N^{-1} \cdot L_{x_0} \xi \rangle}{\langle L_{x_0} \xi, L_{x_0} \xi \rangle} \cdot \frac{\langle L_{x_0} \xi, L_{x_0} \xi \rangle}{\langle \xi, \xi \rangle} \\ &\geq \sigma_{\min}(L_{x_0})^2 \cdot \min_{0 \neq v} \frac{\langle v, N^{-1} \cdot v \rangle}{\langle v, v \rangle} = \sigma_{\min}(L_{x_0})^2 \cdot \lambda_{\min}(N^{-1}) \\ &= \sigma_{\min}(L_{x_0})^2 \cdot \lambda_{\max}(N)^{-1}, \end{aligned}$$

where $\sigma_{\min}(L_{x_0})$ denotes the smallest singular value of L_{x_0} . This yields

$$\lambda_{\max}((L_{x_0}^\top N^{-1} L_{x_0})^{-1}) \leq \sigma_{\min}(L_{x_0})^{-2} \cdot \lambda_{\max}(N) \xrightarrow{N \rightarrow 0} 0.$$

This now implies $(L_{x_0}^\top N^{-1} L_{x_0})^{-1} \rightarrow 0$.

By, e.g., applying the Courant-Fischer Min-Max theorem (see, e.g., [174], Theorem 1.9) with the spaces $\mathbb{R} \cdot \eta$ and $\{\eta\}^\perp$, it follows that once $(L_{x_0}^\top N^{-1} L_{x_0})^{-1}$ is sufficiently small, the matrix

$(L_{x_0}^\top N^{-1} L_{x_0})^{-1} + \eta \cdot \eta^\top$ has two eigenvalues close to 0 and one eigenvalue close to $\|\eta\|^2$. Applying [174], Theorem 3.9 (with $\tilde{u} := \frac{M_0}{\|M_0\|}$) then yields $M_{\text{UG}} \rightarrow M_0$. This finishes the proof of 1).

To prove 2), note that, under the assumptions of 2), we have

$$M_{\text{UG},\rho} = \nu_{\max} \left((L_{x_0}^\top N_1^{-1} L_{x_0})^{-1} + \frac{\eta \cdot \eta^\top}{\rho} \right)$$

As $\rho \rightarrow \infty$, we now have $\frac{\eta \cdot \eta^\top}{\rho} \rightarrow 0$, and we can copy the proof strategy from 1). \square

Theorem 3.3.17 describes the qualitative behaviour of the UG orientation reconstruction process in noisy single-source scenarios. One can, in fact, go much further than that. If you look at the proof of theorem 3.3.17 above, you see that the UG orientation estimate is given by a rank 1 update of the matrix $L_{x_0}^\top N^{-1} L_{x_0}$. Some readers may already be aware that in such a scenario, one can prove quite strong results about the relationship of the eigendecomposition of the original matrix and the matrix after the rank one update, as, e.g., shown in the work by Bunch, Nielsen, and Sorensen [29]. This can be used to give a more quantitative version of the behaviour of the UG orientation estimate. Using this strategy, one can, e.g., investigate to what degree the UG beamformer is able to reconstruct tangential orientation components, similar to the discussions surrounding theorem 3.3.15. One can also discuss edge cases that are left open by theorem 3.3.17, such as what the limiting behaviour in case 2) is if the minimal eigenvalue of $L_{x_0}^\top N_1^{-1} L_{x_0}$ is not simple. We refer interested readers to [30], where we investigated these questions. For the purpose of this thesis, we will finish our investigations into beamforming approaches with theorem 3.3.17.

Chapter 4

Source reconstruction and tDCS stimulation of an epilepsy patient

In the preceding sections, we have been primarily focused on methodological aspects of bio-electromagnetism. In this section, we want to showcase an application of these methods. This application is embedded in the *PerEpi project*.¹ The goal of this project was the development and application of sophisticated methodologies for the purpose of aiding the diagnosis and therapy of refractory focal epilepsy.

Epilepsy is a neurological disease characterized by recurring unprovoked seizures [16]. A lifetime prevalence of epilepsy of 7.6 per 1000 persons (95% CI 6.17-9.38) has been reported [63]. One speaks of a *focal epilepsy* if the electric activity accompanying the seizures is localized to a subset of the brain [195]. The electrical activity can be localized to a single region, in which case one speaks of an unifocal disorder, or to multiple regions, in which case one speaks of a multifocal disorder, or to a hemisphere [180]. Focal epilepsies constitute a significant portion of all epilepsies. In [193] the authors, e.g., report that out of 2200 investigated patients, they observed a focal epilepsy in 1369 cases, i.e., in ca. 62% of the patients. Furthermore, one speaks of a *refractory*, or *drug resistant*, epilepsy if, after trying two appropriate antiepileptic drugs, the patient does not achieve seizure freedom [109]. As ca. 30% of focal epilepsy patients are classified as refractory [164], we thus see that the subset of refractory focal epilepsies, which are in the focus of the PerEpi project, accounts for a large portion of epilepsy cases.

In the case of refractory focal epilepsy, it has been recommended to consider epilepsy surgery, as it has been shown that, when compared to trying further antiepileptic drugs, for certain types of focal epilepsies, a surgery has a much higher chance of leading to seizure freedom, while being both cost-effective and having a low complication rate [232, 230, 50, 155, 228]. Note, however, that, due to a multitude of reasons, surgery is not always an option. Before considering a surgery, one e.g. needs to estimate the so-called *epileptogenic zone*, which is the subset of the cortex that needs to be removed to achieve seizure freedom [167]. This is, in general, a difficult problem without a one-size-fits-all solution. And even if a reasonable estimate of the epileptogenic zone is available, factors such as multifocality [10] or closeness to eloquent cortex [167] might lead to the assessment that the risk of an invasive brain surgery is not justified. Furthermore, even if a surgery is performed, seizure freedom is not guaranteed. In [43], it is, e.g., estimated that

¹<https://www.medizin.uni-muenster.de/perepi/startseite.html>

the probability of remaining seizure free for at least 10 years after the epilepsy surgery is ca. 50%.

If antiepileptic drugs are ineffective, and epilepsy surgery is not an option, a further possibility is given by brain stimulation. Multiple forms of brain stimulation have been shown to have a high probability of reducing the seizure frequency of an epilepsy patient, including vagus nerve stimulation [49], deep brain stimulation [175], and transcranial direct current stimulation (tDCS) [241]. Another emerging stimulation strategy, called EASEE, consists of placing stimulation electrodes directly on the skull of the patient, and has also recently been shown to achieve a seizure frequency reduction in a large proportion of patients [188]. It is, however, important to note that brain stimulation in the majority of cases does not lead to seizure freedom, but “only” to a reduction in seizure frequency. This is a crucial distinction, as in many cases the quality of life of an epilepsy patient is not significantly improved by a reduction in seizure frequency, but only by achieving complete seizure freedom [22].

In the PerEpi project, we contributed to the preparation of epilepsy surgeries. As noted above, a brain resection is ideally based on an accurate localization of the epileptogenic zone. EEG and MEG source reconstruction can contribute to this localization. Concretely, it has been demonstrated both in the EEG case [26] and in the MEG case [52, 164] that an overlap of resection area and source reconstruction is strongly associated with whether seizure freedom is achieved after epilepsy surgery. In [26], it is in particular demonstrated that this association becomes stronger if individualized head models, instead of template ones, are used. This suggests that the integration of individualized and realistic head models in combination with well-understood source reconstruction approaches into the presurgical evaluation of a patient with refractory focal epilepsy might enable a better estimation of the epileptogenic zone, which might in turn increase the proportion of epilepsy patients eligible for surgery, and also the proportion of patients that remain seizure free after surgery. It is, however, also important to curb undue optimism. Even ignoring practical factors such as noisy measurements, inaccurate co-registrations, or sparse sensor coverage, the EEG and MEG inverse problems are fundamentally ill-posed. If Ω is the head domain, and $U \subset \Omega$ is open, it is not difficult to construct primary current densities j^P such that $j^P \neq 0$, $\text{supp}(j^P) \subset U$, and $E = B = 0$ on $\mathbb{R}^3 \setminus U$, where E and B denote the electric and magnetic fields, see, e.g., [83]. Speaking more plainly, there are non-zero primary current densities that cannot be detected by either EEG or MEG. Thus, while EEG and MEG can certainly yield valuable contributions in refractory focal epilepsy, these modalities should be seen as single pieces of a larger puzzle, and have to be complemented by additional diagnostic procedures.

In addition to the presurgical evaluation of epilepsy patients, the aim of PerEpi was also to contribute to brain stimulation for epilepsy patients. Concretely, PerEpi had a focus on individualized tDCS brain stimulation.

The motivation for this was that tDCS, compared to other stimulation approaches, such as vagus nerve stimulation or deep brain stimulation, does not require surgery. tDCS is thus much less invasive than other stimulation modalities, and can avoid some of the side effects that oftentimes accompany the more invasive stimulation approaches [104]. Furthermore, it has been demonstrated that tDCS can be considered as a safe stimulation modality with few, and at most moderate, side effects, at least as long as one stays within the parameters of “conventional tDCS”, i.e., the applied current does not exceed 4 mA and is not applied for longer than one hour per day [4]. Note, however, that the price for this non-invasiveness is that any current aimed at the brain first has to pass through the insulating skull, which, due to the scalp acting as a shunt, leads to less strong and less focal electric fields in the brain when compared to stimulation electrodes

implanted inside the skull [214, 121].

Now, conventional neurophysiological wisdom indicates that the cortex below an anode is *excited*, while the cortex below a cathode is *inhibited* [139]. The natural approach to tDCS stimulation in focal epilepsy is thus to estimate the epileptogenic zone and place a cathode over it. This strategy is called *cathodal stimulation* in the literature, and has indeed been investigated and shown to lead to a significant reduction of seizure frequency and/or interictal epileptiform discharge (IED)² frequency in multiple studies [104, 241, 166]. There are, however, also studies where cathodal stimulation did not achieve a significant reduction of epileptic activity [8, 211].

One reason for the sporadic failure of cathodal tDCS stimulation might be an insufficient personalization of the stimulation montage. There are already works from the 1960s showing that the effect of a tDCS stimulation is associated with the normal component of the current density relative to the cortical surface [35, 21]. As neural sources are oriented normal to the cortical surface, this association can alternatively be framed as saying that the effect of a tDCS stimulation is related to whether the current density is parallel or antiparallel to the targeted neural sources. In more recent studies, it has been further demonstrated that the effect of a current on neural sources depends on the alignment between the two, with better alignments leading to larger stimulation effects [163, 134, 107]. From this point of view it seems natural to not only take the position of a neural target into account, but also its orientation, and set up stimulation montages in such a way that the injected current aligns as good as possible with this orientation. There are, in fact, a number of studies where exactly this strategy is investigated, e.g., [45, 60, 224]. In the PerEpi project, a particular focus was on the development and evaluation of the so-called *distributed constrained maximum intensity (D-CMI)* approach, which was first introduced in [106], with the goal of minimizing the skin sensation during stimulation by distributing the current density over multiple electrodes, while maintaining a good alignment between the injected current and the targeted neural sources.

This chapter is now organized as follows. In section 4.1, we will present a detailed look at the D-CMI algorithm and discuss the strategy we used for stimulation montage optimization in this thesis. Following this, we will focus on the epilepsy patient who was investigated and stimulated in this thesis as part of the PerEpi project. In section 4.2, we will describe the head model creation and source reconstruction of the patient. In section 4.3, we will then show the outcome of the montage optimization and present the result of the tDCS stimulation.

This investigation is part of the larger PerEpi project, and is the fourth time this process has been performed. The first case has been published in [5] and the second has been presented on a poster available here³ and in the dissertation [55] available here.⁴ For the third patient, which was stimulated almost concurrently with the fourth patient we present in this thesis, a source reconstruction and tDCS stimulation montage optimization was performed in the PhD thesis [55], in chapters 8 and 9. The stimulation of the third patient was, however, only carried out after completion of [55], and, at the point of writing this thesis, the data measured during this stimulation is still being processed by our clinical partners.

²IEDs are certain types of sharp transient waves present in the EEG or MEG of the majority of epilepsy patients, see [102]. Here, “interictal” denotes the time period outside of a seizure, with “ictal” in contrast describing the time period during a seizure. IEDs are frequently used to diagnose epilepsy [229].

³https://www.sci.utah.edu/~wolters/PaperWolters/2024/Erdbuegger_Biomag2024.pdf

⁴https://www.sci.utah.edu/~wolters/PaperWolters/2024/Erdbuegger_Dissertation.pdf

4.1 Stimulation montage optimization

As the reader probably expects, there is a lot of variance in how precisely the idea of optimizing the alignment between injected current and targeted neural sources is turned into a computational algorithm. After discretization, the general structure is, however, almost always as follows. First, one specifies a target position, say $x \in \mathbb{R}^3$, and a target orientation, say $o \in \mathbb{R}^3$. If $I \in \mathbb{R}^L$ then denotes a current injection pattern, and j_I the corresponding current density in the head, one tries to maximize the expression

$$\langle j_I(x), o \rangle - P(I),$$

over the set of current injection patterns I , subject to a set of side constraints, where P denotes some penalty term. The expression $\langle j_I(x), o \rangle$ is typically called the *directionality* of the current injection pattern I , and we see that the primary goal of the montage optimization is maximizing this directionality. In the introduction, we already mentioned that in the PerEpi project, and thus also in particular in this thesis, the montage optimization uses the so-called D-CMI approach, which was originally developed in [106], as a starting point. In [106], this algorithm is described as follows. First, choose a penalty parameter $\lambda \geq 0$. The optimized current injection pattern $I_{\text{stim}}^\lambda \in \mathbb{R}^L$ is then given by

$$I_{\text{stim}}^\lambda := \arg \max_{I \in \mathbb{R}^L} \langle j_I(x), o \rangle - \lambda \cdot \|I\|_2 \quad \text{subject to} \quad \begin{cases} \sum_{l=1}^L I_l = 0 \\ \|I\|_1 \leq 2 \cdot C_{\text{total}} \\ \|I\|_\infty \leq C_{\text{max}} \end{cases} \quad (4.1)$$

Here, $C_{\text{total}} > 0$ denotes the limit on the total amount of current injected into the head, which might, e.g., be 4 mA, and $C_{\text{max}} > 0$ denotes the maximal amount of current injected or extracted per electrode, which might, e.g., be 2 mA. Furthermore, note that the first constraint ensures that the amount of current injected into the head is the same as the amount of current extracted from the head. Finally, it is suggested in [106] to choose λ in a personalized manner. To this end, one chooses a threshold $C_0 > 0$, and for a stimulation pattern $I \in \mathbb{R}^L$ defines the l -th electrode to be *active* if $I_l \geq C_0$. Then, one defines a targeted number of active electrodes, say $K \in \mathbb{N}$, and tries to choose λ in such a way that the directionality is not reduced too much while yielding K active electrodes. In practice, this is done by starting at $\lambda = 0$, and increasing λ in small steps as long as the number of active electrodes is at most K .⁵

There are, however, some problems with the D-CMI approach. To start off, it is unclear if the D-CMI approach has even been used in the corresponding publications of our workgroup [106, 105]. Recently, our workgroup published its pipeline to Zenodo, see here⁶ or [54]. If you take a look at “pipeline.zip” in this repository, and open the file “tDCS/dcmi_optimization.m”, you will see that the script does not implement $\|I\|_2$ as a penalty term, but instead implements $\sqrt{\sum_{l=2}^L I_l^2} + |I_1|$ as a penalty term. However, as the authors of [106] and [105] did not publish the scripts used to generate their results, it cannot be known whether the scripts from the pipeline with the incorrect implementation were used, or if the publications utilized a correct implementation that was not shared with the workgroup.

Furthermore, it is claimed in [106] that adding $\|I\|_2$ makes the problem strictly convex, which is

⁵In [106] the authors choose $K = 8$, motivated by the fact that the tDCS device owned by the IBB and used in the study enables the use of up to 8 electrodes. See <https://www.neuroelectrics.com/starstim-8> for a description of the device.

⁶<https://zenodo.org/records/14864487>

not true. Take, e.g., $0 \neq v$ and $0 < \alpha < \beta$. Then

$$\left\| \frac{1}{2}(\alpha \cdot v) + \frac{1}{2}(\beta \cdot v) \right\|_2 = \frac{1}{2} \|\alpha \cdot v\|_2 + \frac{1}{2} \|\beta \cdot v\|_2,$$

i.e., $v \mapsto \|v\|_2$ is not strictly convex. A particular consequence of this is that the optimization problem (4.1) in general does not have a unique maximizer for $\lambda > 0$, contrary to what is claimed in [106]. Another undesirable property of (4.1) is that the maximizer does not depend continuously on the parameter λ .

Theorem 4.1.1. *Assume that there is at least one current injection pattern I fulfilling the side constraints of (4.1) with the property that $\langle j_I(x), o \rangle \neq 0$. Let*

$$\mathcal{D} := \left\{ I \in \mathbb{R}^L \mid \sum_{l=1}^L I_l = 0, \|I\|_1 \leq 2 \cdot C_{\text{total}}, \|I\|_\infty \leq C_{\text{max}} \right\}.$$

Then there exists a $\lambda_0 > 0$ with the following properties.

1. *For each $\lambda < \lambda_0$ we have $I_{\text{stim}}^\lambda \in \mathcal{D} \setminus \text{relint}(\mathcal{D})$, where $\text{relint } \mathcal{D}$ denotes the relative interior of \mathcal{D} .*
2. *For each $\lambda > \lambda_0$ we have $I_{\text{stim}}^\lambda = 0$.*
3. *For $\lambda = \lambda_0$, the optimal injection pattern $I_{\text{stim}}^{\lambda_0}$ is not unique.*

As $0 \in \text{relint}(\mathcal{D})$, we thus in particular see that $\lambda \mapsto I_{\text{stim}}^\lambda$ is discontinuous.

Proof. Let $\mathbb{R}_*^L := \{I \in \mathbb{R}^L \mid \sum_{l=1}^L I_l = 0\}$. Then \mathbb{R}_*^L is the affine hull of \mathcal{D} , and as $0 \in \text{relint}(\mathcal{D})$, there is an $\epsilon > 0$ such that $S_\epsilon(0) \cap \mathbb{R}_*^L \subset \mathcal{D}$, where $S_\epsilon(0)$ denotes the surface of the ball of radius ϵ around 0.

Additionally, the mapping $I \mapsto j_I(x)$ is linear, which directly follows from the definition of the EEG forward problem in 2.4.1. This in particular implies that there is a vector $0 \neq w \in \mathbb{R}_*^L$ such that for all $I \in \mathbb{R}_*^L$ we have

$$\langle j_I(x), o \rangle = \langle w, I \rangle.$$

Now define $\lambda_0 := \|w\|_2$. If $\lambda > \lambda_0$, we have for $0 \neq I \in \mathcal{D}$ that

$$\langle w, I \rangle - \lambda \cdot \|I\|_2 \leq \|w\|_2 \cdot \|I\|_2 - \lambda \cdot \|I\|_2 = (\lambda_0 - \lambda) \cdot \|I\|_2 < 0,$$

and $\langle w, 0 \rangle - \lambda \cdot \|0\| = 0$, i.e., $I_{\text{stim}}^\lambda = 0$.

If instead $0 \leq \lambda < \lambda_0$, we have for $I' := \epsilon \cdot \frac{w}{\|w\|_2}$ that $I' \in \mathcal{D}$, and

$$\langle w, I' \rangle - \lambda \cdot \|I'\|_2 = \epsilon \cdot (\lambda_0 - \lambda) > 0,$$

i.e., the maximal value of (4.1) is strictly positive, and is thus in particular not attained in 0. Let the maximum be attained in $\hat{I} \in \mathcal{D}$. If $\hat{I} \in \text{relint } \mathcal{D}$, there must exist a $\rho > 0$ such that $(1 + \rho) \cdot \hat{I} \in \mathcal{D}$. But then

$$\begin{aligned} \langle w, (1 + \rho)\hat{I} \rangle - \lambda \cdot \|(1 + \rho)\hat{I}\|_2 &= (1 + \rho) \cdot \underbrace{\left(\langle w, \hat{I} \rangle - \lambda \cdot \|\hat{I}\|_2 \right)}_{> 0, \text{ as } \hat{I} \text{ is a maximizer}} \\ &> \langle w, \hat{I} \rangle - \lambda \cdot \|\hat{I}\|_2, \end{aligned}$$

which is a contradiction to \hat{I} being a maximizer of (4.1). Hence, any maximizer is contained in $\mathcal{D} \setminus \text{relint } \mathcal{D}$.

Finally, if $\lambda = \lambda_0$, we have for an arbitrary $I \in \mathcal{D}$ that

$$\langle w, I \rangle - \lambda \cdot \|I\|_2 \leq \|w\|_2 \cdot \|I\|_2 - \lambda \cdot \|I\|_2 = (\lambda_0 - \lambda) \cdot \|I\|_2 = 0,$$

i.e., 0 is a maximizer of (4.1). If, however, $0 < \rho \leq \epsilon$ is arbitrary, we also have $\rho \cdot \frac{w}{\|w\|_2} \in \mathcal{D}$ and

$$\langle w, \rho \cdot \frac{w}{\|w\|_2} \rangle - \lambda \cdot \|\rho \cdot \frac{w}{\|w\|_2}\|_2 = \rho \cdot (\|w\|_2 - \lambda) = 0.$$

Hence $\rho \cdot \frac{w}{\|w\|_2}$ is also a maximizer, and we see that (4.1) has infinitely many maximizers in this case. \square

We thus see that, at a certain value of the parameter λ , the D-CMI optimizers collapse to zero. This is illustrated in figure 4.1. In [106], the individualization of the parameter λ is performed by increasing λ in tiny steps until the corresponding optimizer has the desired structure. But if the optimizers do not depend continuously on λ , even small changes in λ can have a large effect on the optimizers. We think that, for an algorithm that is based on small modifications of the input parameter, this is an unappealing property.

Given the preceding discussion, we argue that the D-CMI approach from (4.1) should be modified as follows. For a penalty parameter $\lambda \geq 0$, we define the *modified D-CMI* current injection pattern $\hat{I}_{\text{stim}}^\lambda \in \mathbb{R}^L$ via

$$\hat{I}_{\text{stim}}^\lambda := \arg \max_{I \in \mathbb{R}^L} \langle j_I(x), o \rangle - \lambda \cdot \|I\|_2^2 \quad \text{subject to} \quad \begin{cases} \sum_{l=1}^L I_l = 0 \\ \|I\|_1 \leq 2 \cdot C_{\text{total}} \\ \|I\|_\infty \leq C_{\text{max}} \end{cases} \quad (4.2)$$

In this case, the penalty function $I \mapsto \|I\|_2^2$ is strictly convex, as can be seen by directly writing out the definition of strict convexity. But then, the functional

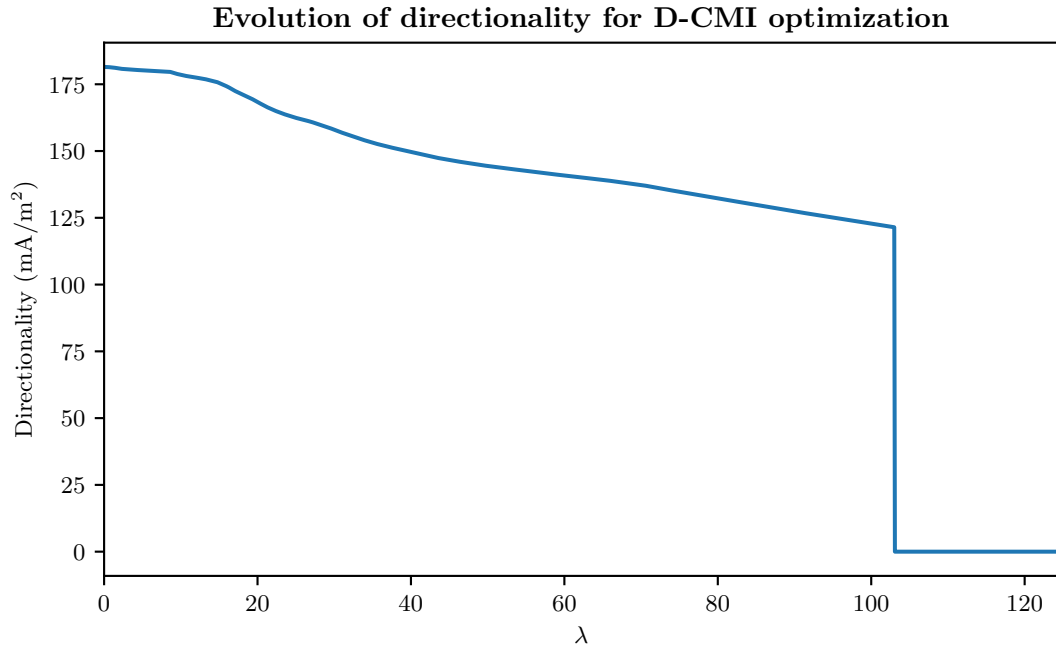
$$f_\lambda : \mathcal{D} \rightarrow \mathbb{R}; I \mapsto \langle j_I(x), o \rangle - \lambda \cdot \|I\|_2^2, \quad (4.3)$$

where \mathcal{D} is defined as in theorem 4.1.1, is strictly concave, which implies that the maximization problem (4.2) has a unique solution for $\lambda > 0$. We will shortly discuss the case $\lambda = 0$ in more detail. For now, we simply note that the $\lambda = 0$ case is typically called the *constrained maximum intensity (CMI)* approach in the literature [106, 60].

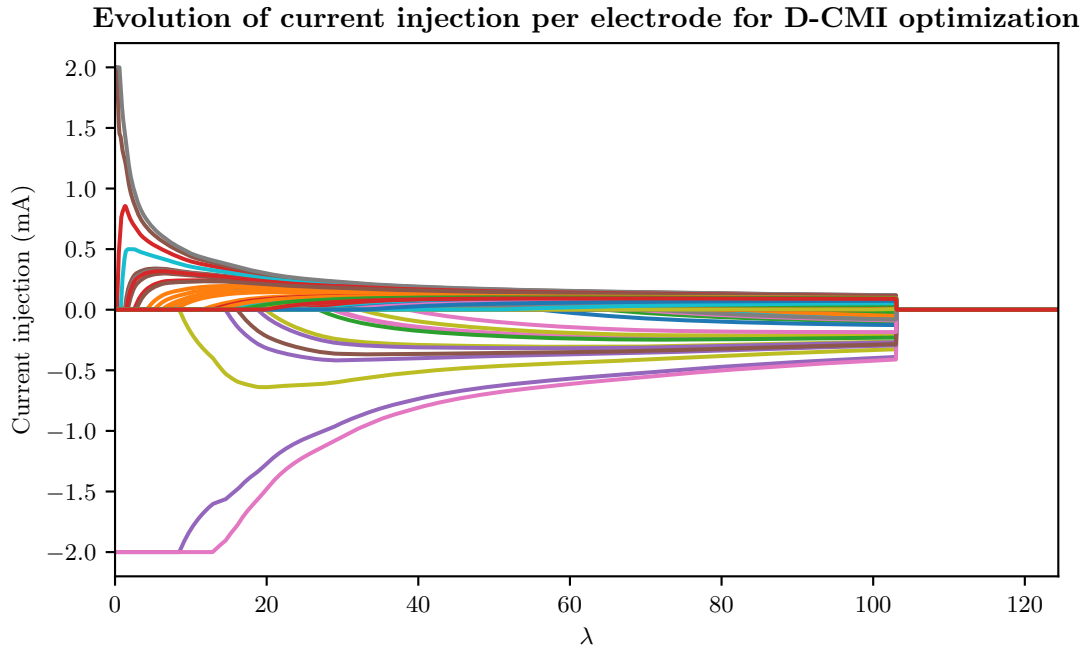
It turns out that the modified D-CMI approach has already been deeply investigated. Concretely, the master's thesis [66] treats (4.2) as the definition of the D-CMI approach.⁷ Thus, all comparisons and investigations in that master's thesis apply to the modified D-CMI approach. We thus refer the reader to [66] for numerous theoretical and numerical investigations, and contend ourselves here with presenting a selection of theoretical properties of (4.2) that differentiate it from the classical D-CMI approach as it is defined in (4.1).

Theorem 4.1.2. *Let \mathcal{D} be defined as in theorem 4.1.1. Furthermore, assume that there is at least one current injection pattern $I \in \mathcal{D}$ with the property that $\langle j_I(x), o \rangle \neq 0$, and for $\lambda > 0$ let $\hat{I}_{\text{stim}}^\lambda$ be defined as in (4.2). We then have the following.*

⁷In [66], see section 2.2, at the bottom of page 56.



(a) Directionality of the optimizing current injection pattern plotted against the value of λ . Notice that the directionality is monotonically decreasing in λ .



(b) The optimizer had 74 electrode positions available for montage optimization. Thus, for each choice of λ , the optimization yields 74 current injection values. In this plot, we visualize for each electrode the optimized injection as a function of λ , i.e., each line corresponds to the current injection of a single electrode.

Figure 4.1: Visualization of a D-CMI optimization result for different values of λ , see (4.1). The target of the optimization was the dipole reconstructed from the EEG data shown in figure 4.10. The optimization used the parameters $L = 74$, $C_{\text{total}} = 4 \text{ mA}$, and $C_{\text{max}} = 2 \text{ mA}$. Notice how the optimizers collapse to zero past a certain value of the parameter λ . The optimizers were computed using the CVX toolbox [37, 76]. To generate the plots, 1500 values for λ evenly spaced between 0 and ca. 125 were sampled.

1. For $\lambda = 0$ the optimization problem (4.2) has in general no unique solution.⁸ There is, however, a unique solution with minimal 2-norm. We will denote by $\hat{I}_{\text{stim}}^0 \in \mathcal{D}$ the unique maximizer of (4.2) with minimal 2-norm.
2. For each $\lambda \geq 0$ we have $\hat{I}_{\text{stim}}^\lambda \neq 0$.
3. The mapping $[0, \infty) \rightarrow \mathcal{D}; \lambda \mapsto \hat{I}_{\text{stim}}^\lambda$ is continuous.

Proof. As in theorem 4.1.1, there exists a $w \in \mathbb{R}_*^L := \{I \in \mathbb{R}^L \mid \sum_{l=1}^L I_l = 0\}$ with the property that for each $I \in \mathbb{R}_*^L$ we have $\langle j_I(x), o \rangle = \langle w, I \rangle$.

We begin with 2). By assumption, there exists $\tilde{I} \in \mathcal{D}$ with $\langle w, \tilde{I} \rangle > 0$. If we define f_λ as in (4.3), we see that $f_\lambda(0) = 0$. For $\lambda = 0$, this immediately yields $\max_{I \in \mathcal{D}} f_0(I) > 0$, which implies that any maximizer of equation (4.2) is non-zero. For $\lambda > 0$, we can compute

$$f_\lambda(t \cdot \tilde{I}) = t \cdot \langle w, \tilde{I} \rangle - t^2 \cdot \left(\lambda \cdot \|\tilde{I}\|_2^2 \right),$$

which as a function of t has two roots at $t = 0$ and $t = \frac{\langle w, \tilde{I} \rangle}{\lambda \|\tilde{I}\|_2^2} > 0$, and is thus positive for a value of t between these roots. This again implies $\max_{I \in \mathcal{D}} f_\lambda(I) > 0$ and thus $\hat{I}_{\text{stim}}^\lambda \neq 0$.

We now first show that $\lambda \mapsto \hat{I}_{\text{stim}}^\lambda$ is continuous on $(0, \infty)$. To this end, let $\lambda_n \rightarrow \lambda \in (0, \infty)$. We need to show that $\hat{I}_{\text{stim}}^{\lambda_n} \rightarrow \hat{I}_{\text{stim}}^\lambda$. If we assume that $\hat{I}_{\text{stim}}^{\lambda_n} \rightarrow \hat{I}_{\text{stim}}^\lambda$ is false, we can find an $\epsilon > 0$ and a subsequence $(\hat{I}_{\text{stim}}^{\lambda_{n_k}})_{k \in \mathbb{N}}$ such that for each $k \in \mathbb{N}$ we have $|\hat{I}_{\text{stim}}^{\lambda_{n_k}} - \hat{I}_{\text{stim}}^\lambda| > \epsilon$. But since \mathcal{D} is compact, there exists a convergent subsequence $\hat{I}_{\text{stim}}^{\lambda_{n_{k_l}}} \rightarrow I' \in \mathcal{D}$. As \mathcal{D} is bounded, it furthermore follows that $f_{\lambda_n} \rightarrow f_\lambda$ uniformly. Together, this yields

$$|f_{\lambda_{n_{k_l}}}(\hat{I}_{\text{stim}}^{\lambda_{n_{k_l}}}) - f_\lambda(I')| \leq \underbrace{|f_{\lambda_{n_{k_l}}}(\hat{I}_{\text{stim}}^{\lambda_{n_{k_l}}}) - f_\lambda(\hat{I}_{\text{stim}}^{\lambda_{n_{k_l}}})|}_{\leq \|f_{\lambda_{n_{k_l}}} - f_\lambda\|_\infty \rightarrow 0} + \underbrace{|f_\lambda(\hat{I}_{\text{stim}}^{\lambda_{n_{k_l}}}) - f_\lambda(I')|}_{\rightarrow 0, \text{ as } f_\lambda \text{ is continuous}} \rightarrow 0,$$

i.e., $f_{\lambda_{n_{k_l}}}(\hat{I}_{\text{stim}}^{\lambda_{n_{k_l}}}) \rightarrow f_\lambda(I')$. This implies for an arbitrary $I \in \mathcal{D}$ that

$$f_\lambda(I) = \lim_{l \rightarrow \infty} f_{\lambda_{n_{k_l}}}(I) \leq \lim_{l \rightarrow \infty} f_{\lambda_{n_{k_l}}}(\hat{I}_{\text{stim}}^{\lambda_{n_{k_l}}}) = f_\lambda(I').$$

Hence I' is a maximizer of f_λ . But as f_λ is strictly concave, it has a unique maximizer, which implies $I' = \hat{I}_{\text{stim}}^\lambda$. But this is a contradiction, as this would imply $\hat{I}_{\text{stim}}^{\lambda_{n_{k_l}}} \rightarrow \hat{I}_{\text{stim}}^\lambda$, which by construction of the subsequence $(\hat{I}_{\text{stim}}^{\lambda_{n_k}})_{k \in \mathbb{N}}$ is impossible. Hence the assumption is wrong, and we have $\hat{I}_{\text{stim}}^{\lambda_n} \rightarrow \hat{I}_{\text{stim}}^\lambda$.

We now investigate 1) and 3) for the $\lambda = 0$ case.

First, we will give an example where the maximizer of (4.2) is not unique. Consider, e.g., the case where $L = 6$, $C_{\text{total}} = 4$, $C_{\text{max}} = 2$, and $w = (1, 1, 1, -1, -1, -1) \in \mathbb{R}_*^6$. Then we have for an arbitrary $I \in \mathcal{D}$ that

$$f_0(I) = \langle w, I \rangle = I_1 + I_2 + I_3 - I_4 - I_5 - I_6 \leq |I_1| + |I_2| + |I_3| + |I_4| + |I_5| + |I_6| = \|I\|_1 \leq 8.$$

⁸Later on, we will, however, see that under very mild conditions there is in fact a unique solution in the $\lambda = 0$, i.e., the CMI, case.

If we now look at $I^1 = (2, 2, 0, -2, -2, 0)$ and $I^2 = (2, 0, 2, -2, 0, -2)$, we have $I^1, I^2 \in \mathcal{D}$, and furthermore

$$f_0(I^1) = f_0(I^2) = 8.$$

Hence, both I^1 and I^2 solve the maximization problem (4.2) and we see that the CMI approach does in general not have a unique solution.

We now return to the general case. In the proof of 2), it was shown that $\max_{I \in \mathcal{D}} f_0(I) > 0$. But as f_0 is linear, it follows that $\mathcal{D}' := \mathcal{D} \cap \{I \in \mathbb{R}^L \mid f_0(I) = \max_{J \in \mathcal{D}} f_0(J)\}$, i.e., the set of maximizers of (4.2), is a non-empty, compact, and convex subset of \mathbb{R}^L that does not contain 0. As $\|\cdot\|_2$ is induced by an inner product, it follows from standard theory, e.g., by theorem 12.3 from [170], that \mathcal{D}' contains a unique element of minimal 2-norm. In the following, we denote this element by $\hat{I}_{\text{stim}}^0 \in \mathcal{D}$.

To finish the proof, it only remains to show that for a sequence $(\lambda_n)_{n \in \mathbb{N}}$ in $(0, \infty)$ with $\lambda_n \rightarrow 0$ we always have $\hat{I}_{\text{stim}}^{\lambda_n} \rightarrow \hat{I}_{\text{stim}}^0$. Assume that this is not the case. As in the proof of the $\lambda > 0$ case, we can then find an $\epsilon > 0$ and a subsequence $(\lambda_{n_k})_{k \in \mathbb{N}}$ such that for each $k \in \mathbb{N}$ we have $|\hat{I}_{\text{stim}}^{\lambda_{n_k}} - \hat{I}_{\text{stim}}^0| > \epsilon$. Again using the same argument as in the $\lambda > 0$ case, we can then find a convergent subsequence $\hat{I}_{\text{stim}}^{\lambda_{n_{k_l}}} \rightarrow I' \in \mathcal{D}$, and show that I' is a maximizer of f_0 in \mathcal{D} , i.e., $I' \in \mathcal{D}'$. By construction of the subsequence $(\lambda_{n_k})_{k \in \mathbb{N}}$ we have $I' \neq \hat{I}_{\text{stim}}^0$. But by construction of \hat{I}_{stim}^0 , this implies $\|I'\|_2^2 > \|\hat{I}_{\text{stim}}^0\|_2^2$. But then there exists an $l_0 \in \mathbb{N}$ with the property that $\|\hat{I}_{\text{stim}}^{\lambda_{n_{k_{l_0}}}}\|_2^2 > \|\hat{I}_{\text{stim}}^0\|_2^2$. As \hat{I}_{stim}^0 is a maximizer of f_0 in \mathcal{D} , this finally implies

$$f_{\lambda_{n_{k_{l_0}}}}(\hat{I}_{\text{stim}}^{\lambda_{n_{k_{l_0}}}}) = \langle w, \hat{I}_{\text{stim}}^{\lambda_{n_{k_{l_0}}}} \rangle - \lambda_{n_{k_{l_0}}} \cdot \|\hat{I}_{\text{stim}}^{\lambda_{n_{k_{l_0}}}}\|_2^2 < \langle w, \hat{I}_{\text{stim}}^0 \rangle - \lambda_{n_{k_{l_0}}} \cdot \|\hat{I}_{\text{stim}}^0\|_2^2,$$

which is a contradiction to $\hat{I}_{\text{stim}}^{\lambda_{n_{k_{l_0}}}}$ being a maximizer of $f_{\lambda_{n_{k_{l_0}}}}$. Hence the assumption is wrong, and we must have $\hat{I}_{\text{stim}}^{\lambda_n} \rightarrow \hat{I}_{\text{stim}}^0$. This shows that the map $\lambda \rightarrow \hat{I}_{\text{stim}}^{\lambda}$ is also continuous in 0. \square

When compared to the original D-CMI formulation from (4.1), we argue that, due to the continuous dependence of the stimulation montage on the penalty parameter, the modified D-CMI formulation from (4.2) is better suited to the montage personalization approach based on small modifications of the penalty parameter as it is suggested by [106]. Accordingly, we used (4.2) in the computation of the tDCS stimulation montage we will present later on in this thesis.

Furthermore, we believe that (4.2) is more in the spirit of what the authors of [106] originally intended. When introducing the D-CMI approach, they refer to [224]. In [224], the authors always use $\|I\|_2^2$ as a penalty term. Furthermore, if you, e.g., compare equation (3) in [106] with the content of section 3.3 in [224], you will see that both papers claim to talk about the same functional they want to optimize, but while [224] uses $\|I\|_2^2$ in their functional, [106] uses $\|I\|_2$. We thus think that the use of $\|I\|_2$ in (4.1) arises from a misunderstanding of [224], as the references and motivations in [106], in far as we understand them, indicate that the authors intended to use [224] as a motivation, which would, however, lead to the use of $\|I\|_2^2$ as a penalty term.

As a final topic in this subsection, we want to spend some paragraphs talking about the CMI approach, i.e., (4.2) in the special case $\lambda = 0$. Concretely, we first want to present how the CMI optimizers can be computed in an extremely simple way, without needing any advanced optimization toolbox. Secondly, we want to show how the CMI optimizers can be interpreted in a remarkably elegant way using the EEG/tDCS reciprocity laws from subsection 2.9.2. These

results can also be found in appendix B of [60]. We will, however, present a somewhat different argument than [60], which we argue is both simpler and conceptually clearer.

Theorem 4.1.3. *We want to explicitly compute a CMI stimulation montage, i.e., a maximizer of (4.2) for $\lambda = 0$. Let $\mathbb{R}_*^L := \{I \in \mathbb{R}^L \mid \sum_{l=1}^L I_l = 0\}$. Then there exists a unique $w \in \mathbb{R}_*^L$ such that for all $I \in \mathbb{R}_*^L$ we have*

$$\langle j_I(x), o \rangle = \langle w, I \rangle.$$

After reordering the electrodes, we can WLOG assume $w_1 \geq w_2 \geq \dots \geq w_{L-1} \geq w_L$. Let $C_{\text{total}} > 0$ and $C_{\text{max}} > 0$ denote the side constraint parameters from (4.2). Then, there exists a unique $k \in \mathbb{N}_0$ and $0 \leq t < C_{\text{max}}$ such that $C_{\text{total}} = k \cdot C_{\text{max}} + t$. Furthermore, we assume that $L \geq 2 \cdot (k + 1)$. Then, the montage

$$I^{\text{CMI}} := (\underbrace{C_{\text{max}}, \dots, C_{\text{max}}}_{k \text{ times}}, t, 0, \dots, 0, -t, \underbrace{-C_{\text{max}}, \dots, -C_{\text{max}}}_{k \text{ times}})$$

is a CMI maximizer. If $w_k > w_{k+1} > w_{k+2}$ and $w_{L-k-1} > w_{L-k} > w_{L-k+1}$, the montage I^{CMI} is the unique CMI maximizer.

Proof. As the mapping $f_0 : \mathbb{R}_*^L \rightarrow \mathbb{R}; I \mapsto \langle j_I(x), o \rangle$ is linear, and \mathbb{R}_*^L is a Hilbert space, the existence of a unique w as claimed in the theorem is a direct consequence of the Riesz representation theorem.⁹

Note that $\sum_{l=1}^L I_l^{\text{CMI}} = 0$, $\|I^{\text{CMI}}\|_1 = 2 \cdot (k \cdot C_{\text{max}} + t) = 2 \cdot C_{\text{total}}$, and $\|I^{\text{CMI}}\|_\infty = C_{\text{max}}$, i.e., we have $I^{\text{CMI}} \in \mathcal{D}$.

Now let $I \in \mathcal{D}$ be arbitrary. Our goal is to show that $f_0(I) \leq f_0(I^{\text{CMI}})$. As $f_0(0) = 0 \leq f_0(I^{\text{CMI}})$, we can assume $I \neq 0$. Then, let

$$P(I) := \{i \in \{1, \dots, L\} \mid I_i > 0\}$$

and

$$N(I) := \{i \in \{1, \dots, L\} \mid I_i < 0\}$$

As $I \neq 0$ and $\sum_{l=1}^L I_l = 0$, it follows that both $P(I)$ as well as $N(I)$ are non-empty. Let $\{p_1, \dots, p_K\} := P(I)$ and be $\{n_1, \dots, n_O\} := N(I)$ be enumerations of $P(I)$ and $N(I)$ with the property that $w_{p_1} \geq w_{p_2} \geq \dots \geq w_{p_K}$ and $w_{n_1} \leq w_{n_2} \leq \dots \leq w_{n_O}$. Furthermore, let $J_1 := I_{p_1}, \dots, J_K := I_{p_K}$ and $H_1 := I_{n_1}, \dots, H_O := I_{n_O}$. Then

$$f_0(I) = \sum_{i=1}^K J_i \cdot w_{p_i} + \sum_{l=1}^O H_l \cdot w_{n_l},$$

$J_1, \dots, J_K > 0$, and $H_1, \dots, H_O < 0$. We will investigate the two sums separately. Now, there exists $M \in \mathbb{N}_0$ and $0 \leq s < C_{\text{max}}$ such that $\sum_{i=1}^K J_i = M \cdot C_{\text{max}} + s$. As $\|I\|_1 \leq 2 \cdot C_{\text{total}}$, we must have $M \leq k$. Furthermore, as $\|I\|_\infty \leq C_{\text{max}}$, there exists a sequence $1 \leq i_1 < \dots < i_M \leq K$ such that i_r is the minimal index such that $\sum_{i=1}^{i_r} J_i \geq r \cdot C_{\text{max}}$. But then we can, for $1 \leq r \leq M$, find a $\gamma_r \in [0, 1]$ with the property that

$$\sum_{i=1}^{i_r-1} J_i + \gamma_r \cdot J_{i_r} = r \cdot C_{\text{max}}.$$

⁹Note that we have implicitly used this in the proofs of theorems 4.1.1 and 4.1.2.

This yields

$$\begin{aligned}
\sum_{i=1}^K J_i \cdot w_{p_i} &= \left(\sum_{i=1}^{i_1-1} J_i \cdot w_{p_i} + \gamma_1 \cdot J_{i_1} \cdot w_{p_{i_1}} \right) \\
&\quad + \sum_{r=1}^{M-1} \left((1 - \gamma_r) \cdot J_{i_r} \cdot w_{p_{i_r}} + \sum_{i=i_r+1}^{i_{r+1}-1} J_i \cdot w_{p_i} + \gamma_{r+1} \cdot J_{i_{r+1}} \cdot w_{p_{i_{r+1}}} \right) \\
&\quad + \left((1 - \gamma_M) \cdot J_{i_M} \cdot w_{p_{i_M}} + \sum_{i=i_M+1}^K J_i \cdot w_{p_i} \right) \\
&\leq C_{\max} \cdot w_1 + \sum_{r=1}^{M-1} C_{\max} \cdot w_{r+1} + s \cdot w_{M+1}.
\end{aligned}$$

Notice that $(1 - \gamma_r) \neq 0$ can only happen if $i_r > r$, in which case $p_{i_r} \geq r+1$ and thus $w_{p_{i_r}} \leq w_{r+1}$. As $\sum_{l=1}^L I_l = 0$, we have

$$\sum_{l=1}^O (-H_l) = \sum_{i=1}^K J_i = M \cdot C_{\max} + s.$$

But then we apply the exact same argument as in the treatment of $\sum_{i=1}^K J_i \cdot w_{p_i}$ to the sum $\sum_{l=1}^O H_l \cdot w_{n_l} = \sum_{l=1}^O (-H_l) \cdot (-w_{n_l})$, which yields

$$\sum_{l=1}^O H_l \cdot w_{n_l} \leq C_{\max} \cdot (-w_L) + \sum_{r=1}^{M-1} C_{\max} \cdot (-w_{L-r}) + s \cdot (-w_{L-M}),$$

which, in combination with the previous result, shows that

$$\begin{aligned}
f_0(I) &\leq C_{\max} \cdot \left(\sum_{i=1}^M (w_i - w_{L-i+1}) \right) + s \cdot (w_{M+1} - w_{L-M}) \\
&\leq C_{\max} \cdot \left(\sum_{i=1}^k (w_i - w_{L-i+1}) \right) + t \cdot (w_{k+1} - w_{L-k}) = f_0(I^{\text{CMI}}),
\end{aligned}$$

where the inequality follows from $M \cdot C_{\max} + s = \frac{\|I\|_1}{2} \leq C_{\text{total}} = k \cdot C_{\max} + t$. This shows that I^{CMI} is a maximizer of the CMI functional f_0 . If we additionally assume $w_k > w_{k+1} > w_{k+2}$ and $w_{L-k-1} > w_{L-k} > w_{L-k+1}$, a close look at the derivation above shows that the derived inequality is sharp for $I \neq I^{\text{CMI}}$, i.e., that under this additional assumption I^{CMI} is the unique maximizer. \square

We thus see that, once we know the vector $w \in \mathbb{R}_*^L$ from theorem 4.1.3, it is extremely easy to compute the CMI maximizer, by just “filling up” the montage symmetrically from the extreme entries of w . Furthermore, note that once we know w , we also know the critical value λ_0 at which the classical D-CMI maximizers collapse to zero, as can be seen from the proof of theorem 4.1.1. Additionally, although we did not include the corresponding argument in theorem 4.1.2, it is easy to see that, for sufficiently large $\lambda > 0$, the modified D-CMI maximizer from equation (4.2) is given by $\hat{I}_{\text{stim}}^\lambda = \frac{w}{2\lambda}$.

Now, the “naive” way of computing w consists of numerically solving the tDCS forward problem, e.g., by using a FEM approach based on 2.4.1 or 2.4.2, for $L - 1$ pairs of electrodes. From these

solutions, we can then assemble a matrix $A \in \mathbb{R}^{3 \times L}$ such that $j_I(x) = A \cdot I$, which in turn yields w via $w = (\text{Id}_3 - \frac{1 \cdot 1^\top}{3}) \cdot A^\top \cdot o$.

There is, however, also another way to compute w . This is the content of the next theorem.

Theorem 4.1.4. *Assume that we use either a PEM or a CEM model. In both cases, denote by $l_{x,\sigma(x) \cdot o}^{\text{CAR}} \in \mathbb{R}_*^L$ the common average referenced EEG lead field due to a dipole at position x with moment $\sigma(x) \cdot o$. Then, the vector $w \in \mathbb{R}_*^L$ from theorem 4.1.3 is given by*

$$w = -l_{x,\sigma(x) \cdot o}^{\text{CAR}}$$

Concretely, this means that for an arbitrary current injection pattern $I \in \mathbb{R}_*^L$ we have

$$\langle j_I(x), o \rangle = -\langle I, l_{x,\sigma(x) \cdot o}^{\text{CAR}} \rangle.$$

Proof. Let $l_{x,\sigma(x) \cdot o}^{\text{CAR}} = (U_1, \dots, U_L)$. Furthermore, for $2 \leq l \leq L$ let $u_{l,1}$ denote a tDCS forward solution corresponding to using the first electrode as a cathode and the l -th electrode as an anode with a current injection of 1.¹⁰ If now $I \in \mathbb{R}_*^L$ denotes an arbitrary current injection pattern, and u_I a corresponding tDCS solution, it directly follows from the definitions of the tDCS forward problems that

$$u_I = \sum_{l \geq 2} I_l \cdot u_{l,1}.$$

Using the reciprocity theorems 2.9.6 and 2.9.9, we can now compute

$$\begin{aligned} \langle j_I(x), o \rangle &= -\langle \sigma \nabla u_I(x), o \rangle = -\sum_{l \geq 2} I_l \cdot \langle \sigma \nabla u_{l,1}, o \rangle = -\sum_{l \geq 2} I_l \cdot \langle \nabla u_{l,1}, \sigma \cdot o \rangle \\ &= -\sum_{l \geq 2} I_l \cdot (U_l - U_1) = -\sum_{l \geq 1} I_l \cdot U_l \\ &= \langle -l_{x,\sigma(x) \cdot o}^{\text{CAR}}, I \rangle. \end{aligned}$$

As $-l_{x,\sigma(x) \cdot o}^{\text{CAR}} \in \mathbb{R}_*^L$, this shows $w = -l_{x,\sigma(x) \cdot o}^{\text{CAR}}$. □

Using theorem 4.1.4, we finally arrive at an extremely simple algorithm for computing a stimulation montage $I \in \mathcal{D}$ that maximizes the directionality at the target. First, we compute the EEG topography that the target would generate. Then, we place anodes at the most negative electrodes and cathodes at the most positive electrodes, until we have exhausted our “budget” of current we are allowed to inject, i.e., until the currents at the anodes sum up to C_{total} . Furthermore, note that in many cases one is already in possession of a good estimate of the target’s EEG topography, as the target, e.g., might have been defined based on averaged epileptic spikes, or averaged evoked potentials. In this case, one can just directly read off the most negative and most positive electrodes from the measured data, without needing to go through the trouble of building a head model or performing a finite element simulation. If the averaged data has a sufficiently high SNR, one could even make the argument that this might be more accurate than performing a FEM simulation, as the real measurement data “automatically” uses a correct segmentation and conductivity values. We want to emphasize that the conclusions in this paragraph are not new insights, and can also be found in [58] and [60].

¹⁰In the PEM case, a solution is directly given by such a function. In the CEM case, a solution is given by a pair $(u_{l,1}, U_{l,1})$, where we only need the $u_{l,1}$ part in this argument. We refer to definitions 2.4.1 and 2.4.2 for details.

As a final consideration, we want to point out that using a version of the D-CMI approach, be it the original version as defined in (4.1) or the modified version from (4.2), always trades away a portion of directionality for a more spread out distribution of electrodes. Especially in the case that only a few stimulation sessions are performed, it might be beneficial to prioritize achieving a high directionality, as we expect this to maximize the chance of working out a significant effect. Furthermore, it is debatable to what extent the larger spread of electrodes, as it is produced by the D-CMI approach, translates to an increased tolerability of the stimulation. In many situations, the D-CMI approach yields a montage that distributes current at electrodes that are farther removed from the target, while yielding the same montage pattern as the CMI approach for electrodes close to the target. Typically, D-CMI thus transforms a CMI montage with two hotspots into a montage with one hotspot. But if two hotspots exceed the pain tolerance of the stimulation subject, we think it is likely that one hotspot also exceeds the pain tolerance of the subject. This, e.g., happened during the stimulation of the epilepsy patient we will introduce in the next section. There, we used the modified D-CMI approach for montage optimization, resulting in a montage containing one hotspot of two electrodes close to one another, which were each extracting 2 mA out of the head. This exceeded the pain tolerance of the epilepsy patient, leading to an early termination of the stimulation. For this patient, only a combination of applying lidocaine, i.e., a local anesthetic, and scaling down the currents by a factor of $\frac{1}{2}$ resulted in a tolerable stimulation.

4.2 Source reconstruction of an epilepsy patient

The study presented in this section and in section 4.3 was conducted according to the guidelines of the Declaration of Helsinki and approved by the institution’s ethical review boards (Ethik Kommission der Ärztekammer Westfalen-Lippe und der Westfälischen Wilhelms-Universität Münster, 25.05.2021, Ref. No. 2021-290-f-S). The patient gave written informed consent for their data to be used in anonymized form for scientific publications.

We will first quickly discuss the relevant patient history. Then, we will showcase the steps that we performed in the construction of an individualized, calibrated, and anisotropic tetrahedral head model for the patient. Finally, we will perform a source estimation based on epileptic spikes of the patient, and will define a tDCS target based on this source estimation. At the IBB, the patient was assigned the label P1106 for data processing. For the purpose of anonymization, we will also use this label in this thesis.

P1106 is a female epilepsy patient who, at the time of the first stimulation, was 25 years old. The patient has suffered from epilepsy since the age of 10. The seizure symptoms consist of oral, hand, and leg automatisms. MRI measurements of P1106 show at least three instances of abnormal cortical development. First, the MRI shows polymicrogyria along the left basal temporal and occipital areas. Secondly, the MRI shows bilateral subependymal heterotopia. In figure 4.2, slices of a T1 MRI highlighting these brain malformations are depicted. The epilepsy of P1106 has been classified as refractory by the attending epileptologist. At the point of writing, P1106 is under presurgical evaluation for epilepsy surgery. In the spring of 2024, P1106 underwent a combined EEG/MEG measurement and an MRI measurement at the IBB in Münster. The combined EEG/MEG measurement was performed for a total of ca. 55 min, split into 6 separate runs. The first of these runs lasted ca. 13 min, and was accompanied by median nerve stimulation of the right hand, where over the duration of the run a total of 1932 electrical pulses were applied. The remaining 5 runs lasted 500 s each, during which no stimulus was applied. These last 5 runs were then shared with the patient’s attending epileptologists for marking of epileptic activity.

This revealed that P1106 shows interictal epileptic activity in both EEG and MEG. After the combined EEG/MEG measurement had concluded, a measurement of T1-, T2-, and DT-MRI images was performed.

4.2.1 Constructing a volume conductor model

The first step in the creation of an anisotropic and calibrated volume conductor model consists of generating an initial isotropic uncalibrated head model. In a coarse approximation, this process consists of three substeps.

1. Coregistration of the input data.
2. Segmentation of the tissue compartments.
3. Creation of a FEM mesh.

For P1106, we measured both T1 and T2 MRI data. These two modalities offer complementary strengths for the segmentation of tissue compartments in the human brain. While a T1 MRI can be used to differentiate gray matter from white matter, a T2 MRI can be used to distinguish between CSF, skull compacta, skull spongiosa, and scalp [138], which suggests a combination of these two imaging modalities to achieve the best possible segmentation result. But such a combination is typically not immediately possible, as oftentimes the subject’s T1 and T2 images are misaligned, e.g., due to head movement. To properly combine T1 and T2 images, this misalignment needs to be corrected. The process of trying to perform this correction is called *registration*, or *coregistration*. This is, fortunately, a common problem in brain research and, accordingly, has been investigated in numerous publications. There exists, in particular, effective and well-tested coregistration software that is readily available and free of charge [99, 100]. Furthermore, after the MRI images have been aligned and a segmentation has been created, one still needs to create a proper FEM mesh from this segmentation, which is also far from trivial [3].

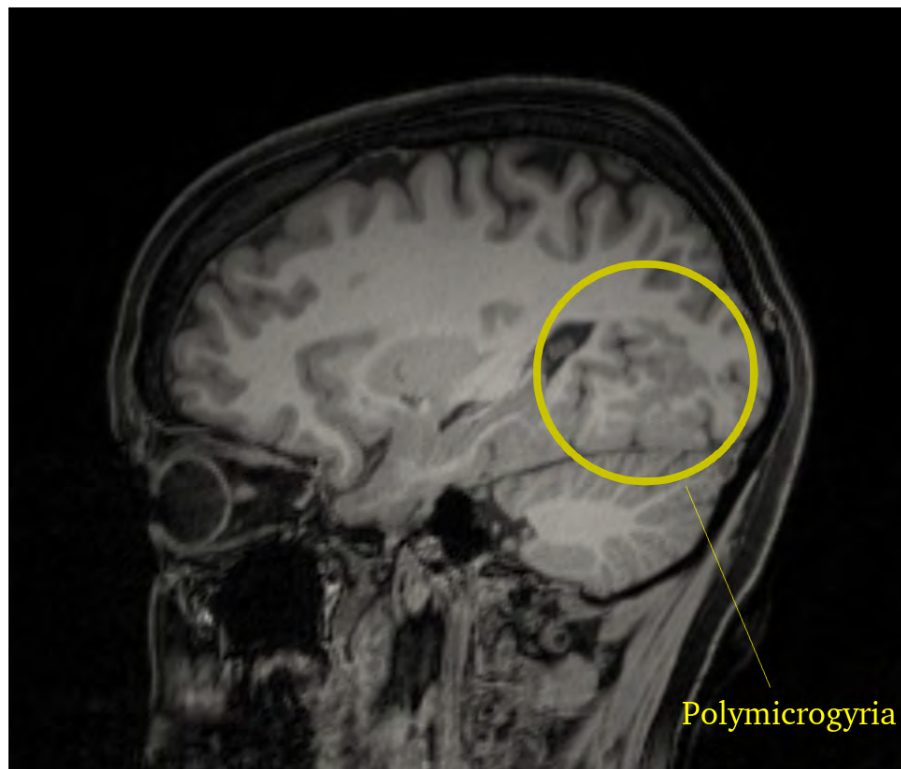
For the creation of an initial head model for P1106, we utilized the CHARM pipeline [158]. CHARM is distributed as part of the SimNIBS toolbox,¹¹ and enables an extremely user-friendly experience for creating head models from MRI images. Concretely, the user only needs to supply a subject ID and paths to the T1 and T2 images, and the CHARM pipeline will automatically perform all the necessary processing to generate a tetrahedral head model, achieving the fabled “button press” solution. In figure 4.3, the final mesh and some intermediary steps towards its creation are visualized. On contemporary hardware at the time of writing, running the CHARM pipeline takes ca. 70 min.

The next step is to associate sensor positions with the FEM mesh. Concretely, we need to bring the electrode positions and MEG magnetometer positions and orientations into the same coordinate system as the FEM mesh. Our strategy to achieve this is based on data measured by a Polhemus digitizer device.¹² This device is used to sample three-dimensional positions on the surface of the subject’s head. The set of sampled positions is chosen in such a way that it contains the electrode positions, as well as the so-called nasion, the left pre-auricular point, and the right pre-auricular point.¹³ Furthermore, the CTF MEG device used by the IBB is equipped with so-called *head localization coils*, which can be used to determine the positions of the MEG magnetometer sensors relative to the head surface [233]. The measurement protocol

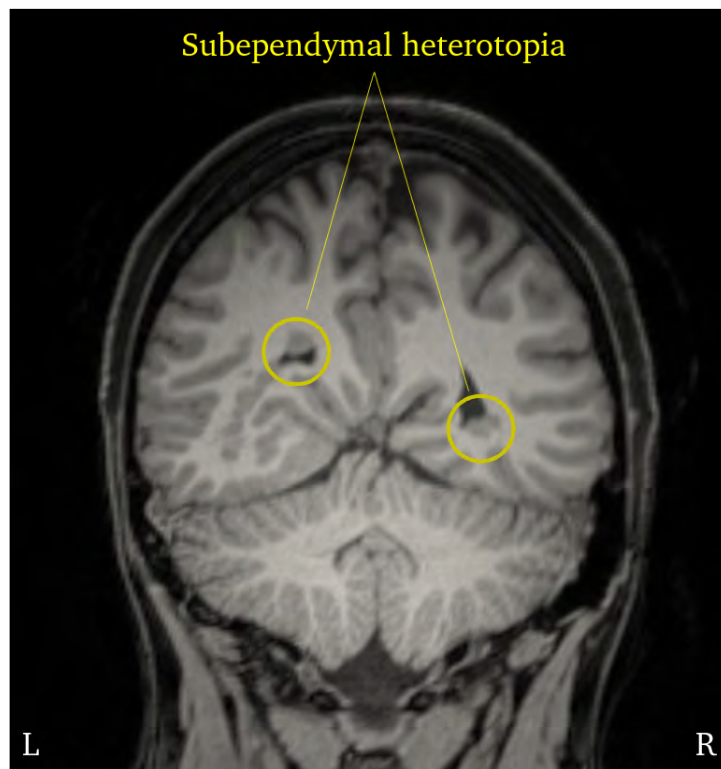
¹¹<https://simnibs.github.io/simnibs/build/html/index.html>

¹²<https://polhemus.com/scanning-digitizing/digitizing-products>

¹³For definitions of the positions see <https://neuroimage.usc.edu/brainstorm/CoordinateSystems> and https://www.fieldtriptoolbox.org/faq/source/anat_landmarks/.



(a) Sagittal slice of a T1 MRI, visualizing the polymicrogyria in the patient's left basal temporal and occipital areas.



(b) Coronal slice of a T1 MRI, visualizing the bilateral subependymal heterotopia in the patient's brain. The letters "L" and "R" denote the left and right hemispheres of the brain.

Figure 4.2: T1 MRI visualizing the abnormal cortical developments of P1106.

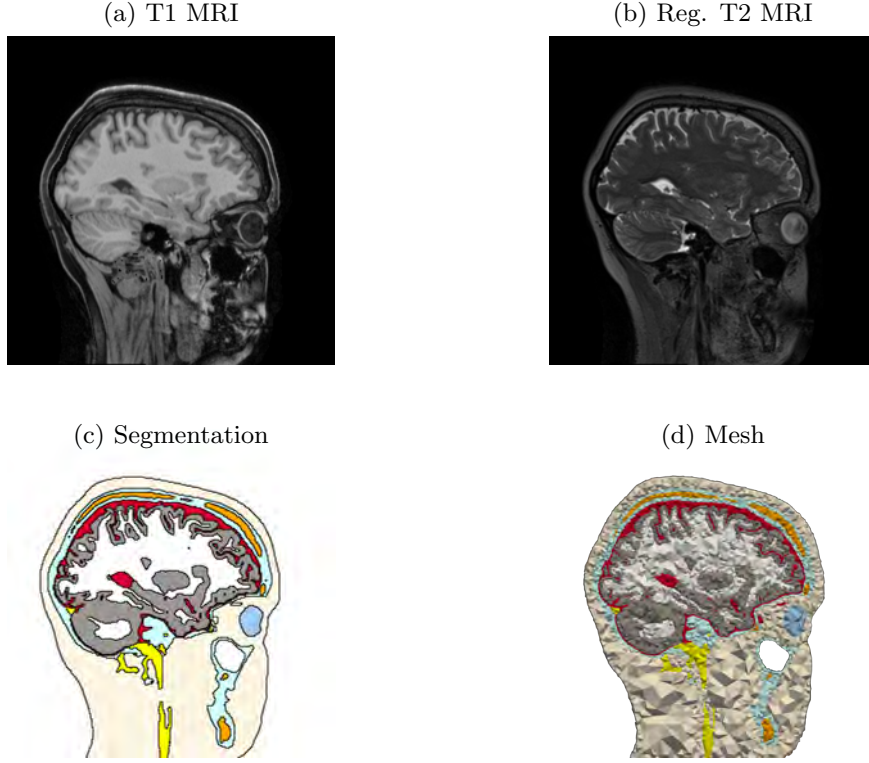


Figure 4.3: Visualization of mesh creation procedure for P1106. The mesh was constructed using the CHARM pipeline [158]. Figure (a) shows a sagittal slice of a T1-MRI of P1106. Figure (b) shows a sagittal slice of a T2-MRI of P1106 that has already been coregistered, i.e., aligned, with the T1-MRI from (a). In figure (c), the segmentation, i.e., the assignment of MRI voxels to tissues, on the sagittal plane from (a) is visualized. In figure (d), each tetrahedron from the final FEM mesh that intersects the sagittal plane from (a) is visualized.

at the IBB automatically utilizes these coils to synchronize the MEG magnetometer coordinate system to the Polhemus surface position coordinate system. Concretely, this means that the data we receive after a measurement already contains the MEG sensors, the EEG sensors, and the Polhemus headshape positions in the same coordinate system, and the task consists “only” of mapping this coordinate system to the coordinate system of the FEM mesh.

To this end, we assume that the head geometry of a subject does not change in the time between the MEG measurement and the MRI measurement. This means that the Polhemus positions and the boundary of the FEM mesh describe the same geometrical object in two potentially different orthonormal coordinate frames. But then there exists a rigid body motion, i.e., a combination of a translation and a rotation, that maps the two coordinate frames to one another. Concretely, if $p_1, \dots, p_K \in \mathbb{R}^3$ denote the coordinates of the Polhemus positions in the MEG coordinate system, there exists a vector $v \in \mathbb{R}^3$ and a rotation $O \in \text{SO}(3)$ such that for $1 \leq i \leq k$ we have that $q_i := v + O \cdot p_i \in \mathbb{R}^3$ describes the coordinates of the i -th Polhemus position in the MRI-derived coordinate system of the FEM head model. In the following, we denote by $T_{v,O}$ the mapping $x \mapsto v + O \cdot x$. We now need some way to estimate v and O .

The classical strategy would be a so-called *landmark-based registration*. Concretely, it is not

difficult to see that the mapping $T_{v,O}$ is determined by its values on three affinely independent points. If we thus identify the nasion, the left pre-auricular point, and the right pre-auricular point on the mesh, we can (in an ideal setting) already estimate v and O from this and use it to transform the sensor positions into the head coordinate system. In our experience, however, the process of manually positioning these landmarks to match the corresponding positions from the Polhemus digitization is subjective, error-prone, and oftentimes leads to inaccurate sensor registrations.

Different members of our workgroup have used different strategies to cope with this. Particularly noteworthy is the strategy implemented in our work group pipeline published in [54],¹⁴ where a significant amount of effort has gone into developing a robust sensor registration strategy for geometry-adapted hexahedral FEM models. A major source of complexity in this implementation originates from the fact that the boundary of a geometry-adapted hexahedral FEM mesh is “rough”, i.e., it contains many bumps.

As the tetrahedral mesh generated by SimNIBS, which we are using for the processing of P1106, has a comparatively smooth boundary, we can get away with a simpler procedure. Concretely, we follow a strategy suggested by the Brainstorm toolbox [200]. Remember that the Polhemus headshape points are sampled on the subject’s head surface. Thus, after applying the “correct” rotation and translation, we expect that the transformed Polhemus positions lie on the boundary of the FEM mesh. If we thus have a candidate coordinate change $T_{v,O}$, we can assess the “goodness” of this transformation by measuring how close the expression

$$f(v, O) := \sum_{i=1}^K \text{dist}(T_{v,O}(p_i), \partial\Omega)^2 \quad (4.4)$$

is to 0, where $\partial\Omega$ denotes the boundary of the FEM mesh. Realistically, we will never achieve $f(v, O) = 0$, both due to measurement inaccuracies as well as to head movement during the Polhemus measurement.¹⁵ We can, however, try to find v and O such that $f(v, O)$ is minimized. To compute v and O , we thus parametrize $\text{SO}(3)$ via Euler angles and apply a Gauss-Newton algorithm to minimize (4.4). For details on the implementation, we refer here.¹⁶

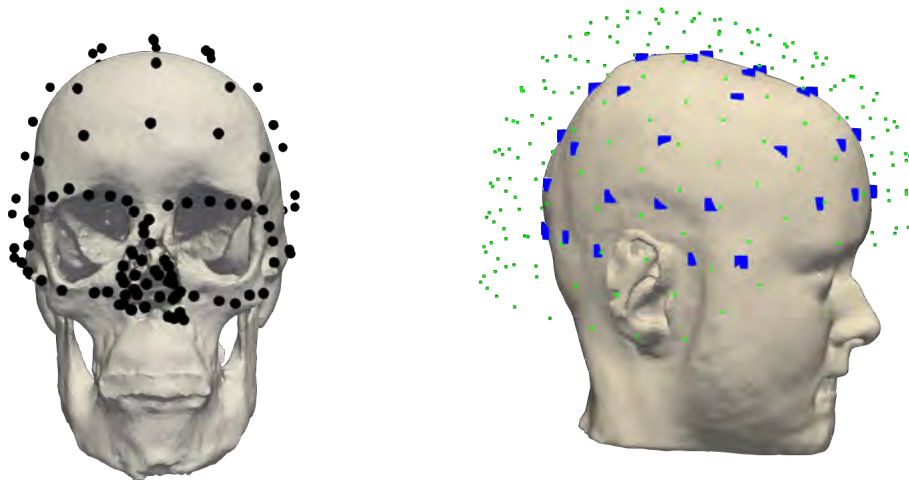
The result of this procedure for P1106 is visualized in figure 4.4. The transformation $T_{v,O}$ computed by applying the Gauss-Newton method to (4.4) is used to transform the EEG as well as the MEG sensor positions. The MEG sensors are then used as-is, but the EEG sensor positions are projected to the closest FEM mesh node on the mesh boundary, i.e., we force the EEG electrode positions to lie on the mesh boundary.

Note that, at this point, we have constructed a model that could be used for FEM forward simulations. There are, however, two more personalization steps we want to perform, which are both related to the choice of conductivity values. If the reader looks at the content of chapter 2 of this thesis, they will see that the conductivity distribution σ inside the head enters every single numerical approach for solving the forward problems, be it EEG, MEG, tDCS, or TMS. The conductivity values, however, will in general vary strongly from person to person, and it has been shown that this variability can result in a considerable uncertainty in both the forward solutions as well as in the inverse source estimations [220, 215]. Furthermore, inside the brain,

¹⁴See the file “Leadfield_generation_calibration/fit_sensors_new.m”

¹⁵During the Polhemus measurement, a pen is pressed with a slight force against the subject’s head at multiple positions. The subject might then slightly move their head in response. Furthermore, typically also the shape of the subject’s nose is sampled, which will oftentimes give in slightly when a force is applied, leading to Polhemus positions *inside* the nose, and not on the nose.

¹⁶https://github.com/brainstorm-tools/brainstorm3/blob/master/toolbox/math/bst_meshfit.m



(a) Final fit of the Polhemus headshape positions after termination of the Gauss-Newton iteration. Each black sphere represents a Polhemus position. As many positions are slightly inside the scalp, we plotted the subject's skull as a reference for better visibility.

(b) Final EEG and MEG sensor fit for P1106. Each blue point represents an electrode position, and each green dot represents an MEG magnetometer. Note that, for better visibility, we did not include the outer shell of MEG magnetometers, the MEG reference sensors, and the magnetometer orientations in the visualization.

Figure 4.4: Result of the sensor fitting algorithm for P1106.

the conductivity σ is in general not a scalar, but an anisotropic tensor, and this distinction can have a crucial impact on both the forward simulation [87, 237] and the inverse source estimation [82]. We want to address both the anisotropy and the conductivity uncertainty.

We start with the anisotropy. The general strategy is to estimate the conductivity tensor σ from DT-MRIs. Note, however, that this is not directly possible, as DT-MRI measures *diffusion tensors*, while we are interested in *conductivity tensors*. We thus need to address two different problems.

1. The estimation of diffusion tensors from DT-MRI.
2. The estimation of conductivity tensors from diffusion tensors.

Problem 1) is a standard task in MRI processing, and can thus be approached with standard toolboxes. A DT image is a collection of multiple MRI images, which describe for each measured voxel the diffusivity in different directions. In the concrete case of P1106, the DT image consisted of 72 individual images. In a first step, these images were brain-extracted and coregistered using FSL [100]. Then, additional susceptibility artifact correction based on the strategy from [172] was performed, using the implementation of the HySCO algorithm in the ACID toolbox available here.¹⁷ Finally, the diffusion tensors were estimated by a simple least squares fit as outlined in [148], using the “dtfit” program from the FSL library.

A link between diffusion tensors and conductivity tensors, i.e., a solution to problem 2), is described in [204] and [87]. In these works, the authors show that under certain modeling

¹⁷<http://www.diffusio.tools.com/index.html>

assumptions, one can derive that the diffusion tensor and the conductivity tensor have the same eigenvectors, and that their eigenvalues are related by an almost linear relationship. If $\sigma(x)$ denotes the conductivity tensor at position x in the brain, and $D(x)$ denotes the diffusion tensor at x , these works thus suggests to model $\sigma(x) = C \cdot D(x)$, where $C > 0$ is some constant.

The remaining task is thus the estimation of this constant C . To this end, we essentially follow the strategy of our work group pipeline from [54], which itself is based on [171]. In this implementation, we only assign anisotropic tensors to white matter elements. The reasoning for this is that while gray matter in adults can be shown to be anisotropic [88], the anisotropy in the gray matter is much weaker than the anisotropy in the white matter, and is difficult to even measure on many MRI machines [194]. We then estimate $C := \frac{\sigma_{\text{WM,default}}}{d_{\text{mean}}}$, where $\sigma_{\text{WM,default}} > 0$ is a standard literature value for isotropic white matter conductivity, and $d_{\text{mean}} > 0$ is the average eigenvalue of the white matter diffusion tensors computed previously. The underlying heuristic behind this choice of C is that it leads to the average white matter conductivity eigenvalue agreeing with the default isotropic white matter conductivity. For the processing of P1106, we followed [158] and selected $\sigma_{\text{WM,default}} := 0.000\,126\,\text{S mm}^{-1}$ as a default isotropic conductivity. As the mean white matter diffusivity eigenvalue for P1106 was $d_{\text{mean}} = 0.000\,846\,2\,\text{mm}^2\text{s}^{-1}$, this resulted in a scaling factor of $C = 0.1489\,\text{S s mm}^{-3}$.¹⁸ The final white matter conductivity tensors for P1106 are visualized in figure 4.5.

The final step towards the construction of a personalized FEM head model is addressing the conductivity uncertainties. While one can find sets of “standard conductivity values” for different brain tissues, e.g., in [158], it is known that these values vary strongly among subjects [128]. Furthermore, it has been demonstrated that a large difference between actual conductivity and assumed conductivity can lead to substantial localization errors [215]. It is thus of great interest to estimate the individual conductivity profile of each subject and adapt the FEM model accordingly.

This is, in fact, a quite common problem in the literature, and different people have suggested different approaches to tackle it, see, e.g., [59] or [70]. We will instead employ a strategy which has been used, and iteratively refined, in our work group at the IBB for more than a decade [11, 6, 185]. This strategy is based on the observation that the magnetic modalities, i.e., MEG and TMS, are in general less influenced by conductivity uncertainties than the electric modalities, i.e., EEG and tES [179, 83].¹⁹ If we thus have a source that can be adequately modeled as a single dipole, as we assumed for the majority of chapter 3, and which can be readily seen in the MEG, we can expect that we can accurately estimate the source position based on this MEG data even in the presence of conductivity uncertainties. But once we have a good estimate of the “true” source position, we can try to adapt the conductivity in such a way that a dipolar source at this position would produce an EEG signal that agrees with the data that was measured, i.e., we search for a conductivity distribution where the simulated EEG topography and the measured EEG topography match as closely as possible. One intricacy with this approach is, however, that it will in general be difficult to define a unique optimal conductivity distribution. In [215], it is, e.g., shown that in the context of reconstructing a single dipole, changing the skull conductivity in one direction can have almost the same effect as changing the scalp conductivity in the opposite direction. As a consequence, one typically only tries to fit *one* conductivity value.

¹⁸Note that based on the literature, e.g., [148], one would expect d_{mean} to be around $0.001\,\text{mm}^2\text{s}^{-1}$, i.e., the value of d_{mean} computed above is in line with what one would expect based on the literature.

¹⁹As an extreme case of this, note that in a multilayer sphere model the magnetic field outside of the model in the MEG case, and the electric field inside the model in the TMS case, are in fact independent of the conductivity profile of the sphere model. For the MEG case, this is shown in [178], and, by the reciprocity laws 2.9.11 and 2.9.12, the corresponding statement for the TMS case is a direct corollary of the MEG case.

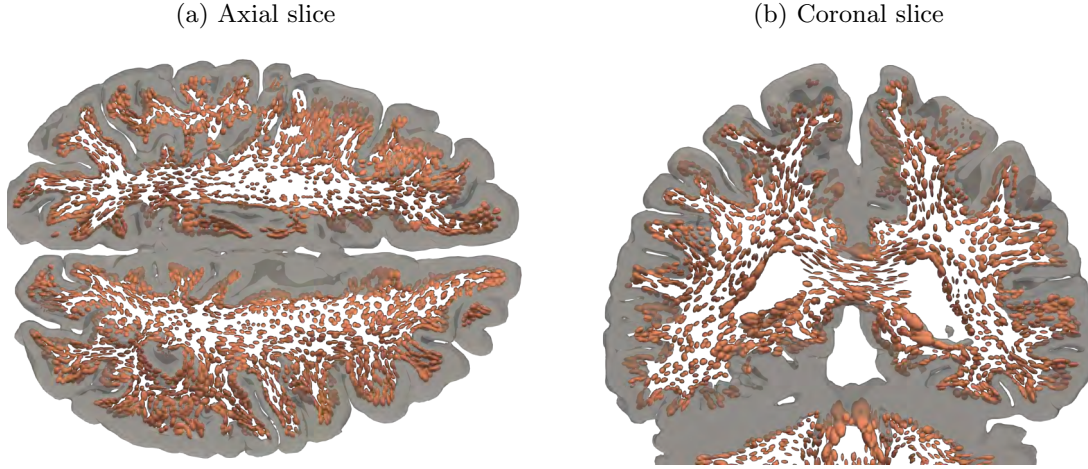


Figure 4.5: Visualization of the DTI-derived anisotropic conductivity tensors computed for the head model of P1106. A tensor is visualized by a red ellipsoid whose principal axes are given by the eigenvectors of the tensor, and whose length along each principal axis corresponds to the eigenvalue of the defining eigenvector. We additionally visualize the gray matter in the slice where the white matter tensors were sampled, to illustrate how the tensors are embedded into the brain. Furthermore, to properly visualize the “direction” of the anisotropy, we only included conductivity tensors with a *fractional anisotropy* of at least 0.3 in the visualization. The fractional anisotropy is a scalar value between 0 and 1, and measures how anisotropic a tensor is, with values closer to 1 indicating a larger anisotropy [148]. The tensor visualizations were performed using ParaView [1].

Concretely, in the IBB pipeline, it was decided to optimize only the skull conductivity, as its value has been shown to have the highest influence on the accuracy of the source estimation [220]. Note, however, that, even if only the skull conductivity is optimized, it has been demonstrated that inaccuracies in the conductivities of other tissues, such as the scalp, can, to some extent, be compensated by an optimized skull conductivity [185].

More concretely, we thus proceed as follows. In the first ca. 13 minutes of the combined EEG/MEG measurement, we perform an electrical stimulation of the subject’s median nerve, where a total of 1932 electrical pulses are applied. The brain activity at ca. 20 ms after applying the electric pulse is typically located on the posterior bank of the central sulcus, with a tangential orientation with respect to the skull [27], which makes it a prime target for MEG source reconstruction [153], and thus also for the conductivity optimization procedure outlined above. For the data preprocessing, we followed the parameter selection of [105]. Concretely, we applied a highpass filter of 20 Hz, a lowpass filter of 250 Hz, and a notch filter at 50 Hz and its harmonics for power line noise. Furthermore, noisy channels and trials were rejected based on visual inspection. The preprocessing was performed in FieldTrip [146]. The average of the filtered trials that were not rejected is shown in figure 4.6. Note how the topographies show a clear tangential dipolar pattern in both EEG and MEG.

SEP/SEF data and topography

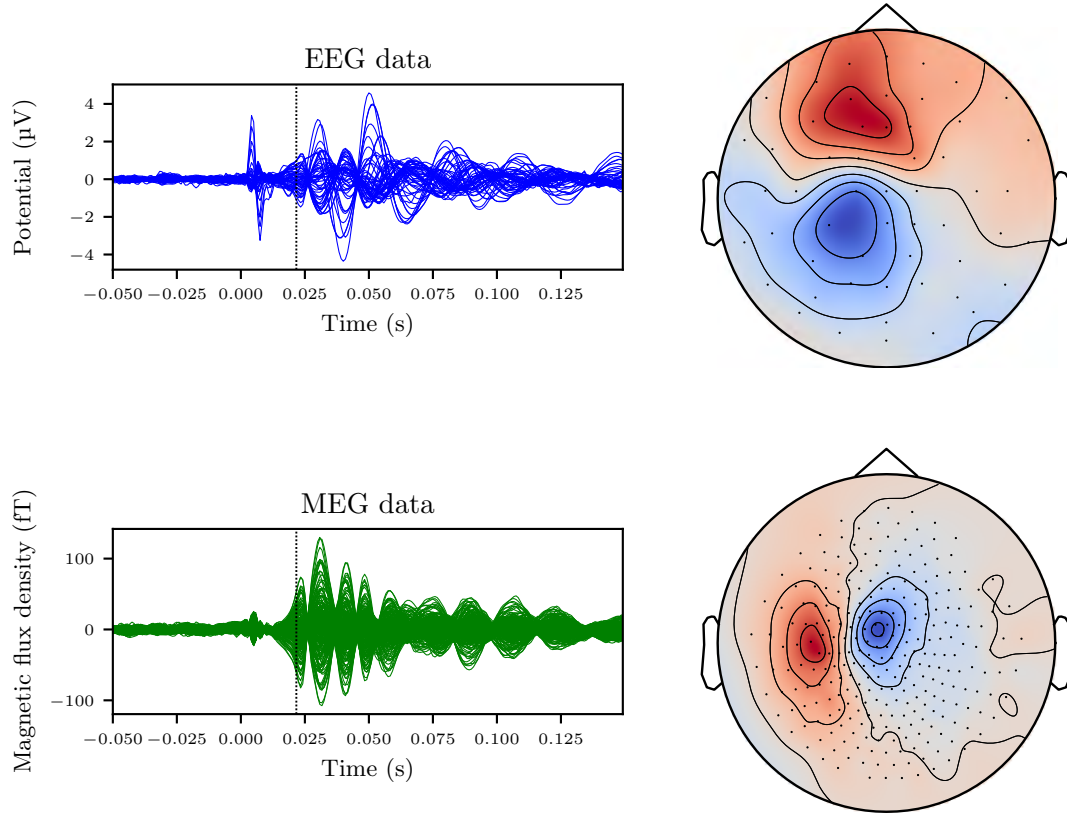


Figure 4.6: Average of EEG and MEG somatosensory evoked activity. In the top row, the EEG data and its topography at 22ms post-stimulus are visualized, and, in the bottom row, the MEG data and its topography at the same timepoint are visualized. The data is visualized as a so-called *butterfly plot*, where the individual signals from each channel are plotted on top of each other. The colormap for the topography is red for positive values and blue for negative values. The dotted line in the data shows the exact time point used for the topography. The mappings of EEG and MEG channel positions to a two-dimensional layout, as well as the head outline, are taken from FieldTrip [146]. Values outside of the layout positions were computed using biharmonic interpolation [177]. To prevent potential confusion, we want to emphasize that, due to our stimulation setup at the IBB, the electrical stimulus takes some time to reach the median nerve. Because of this, our measurements yield slightly later time points than what one would expect based on the literature. Concretely, in a more standard clinical setup, such as, e.g., [27], the topographies visualized above typically occur at 20 ms post-stimulus.

We then constructed a source space with a resolution of 1.5 mm as described in section 2.7.3. Based on this source space, we then computed a MEG lead field using the local subtraction approach as defined in 2.5.17, where this initial computation utilized the default conductivity values from [158]. With this MEG lead field in hand, we then estimated the dipole position using a dipole scan as defined in section 3.1.2, using $C = \text{Id}$. This resulted in a reconstructed dipole

position with a goodness of fit of ca. 96%, as visualized in figure 4.7.

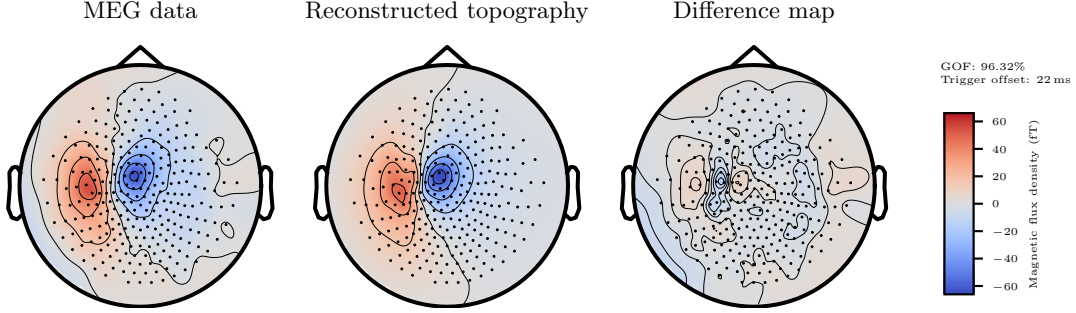


Figure 4.7: Result of a dipole scan applied to the SEF MEG data shown in figure 4.7. The left topography shows the original data, the middle topography shows the lead field of the best-fitting dipole, and the rightmost plot shows the difference between the left and the middle topography.

The general idea is now as follows. Let $x \in \Omega$ denote the position estimated from the MEG dipole scan, and let L denote the number of EEG electrodes. For a scalar $\sigma_{\text{skull}} > 0$, let $L(\sigma_{\text{skull}}) \in \mathbb{R}^{L \times 3}$ denote the EEG lead field at position x under the assumption that

1. the skull conductivity is given by σ_{skull}
2. the white matter conductivity is given by the element-dependent anisotropic tensors estimated from the DTI data
3. the remaining conductivities are set to default values, in our concrete case to the ones from [158].

Furthermore, let $d^{\text{EEG}} \in \mathbb{R}^L$ denote the EEG measurement data at the time point that was used for the MEG dipole scan. We then search for the optimal $\hat{\sigma}_{\text{skull}} > 0$ defined by

$$\hat{\sigma}_{\text{skull}} := \arg \min_{\sigma_{\text{skull}} > 0} \text{rv}_{\text{Id}}(d^{\text{EEG}}, L(\sigma_{\text{skull}})), \quad (4.5)$$

where the residual variance is defined as in equation (3.2). Experience suggests that the map $\sigma_{\text{skull}} \mapsto \text{rv}_{\text{Id}}(d^{\text{EEG}}, L(\sigma_{\text{skull}}))$ is a convex function that assumes a global minimum within the range of one order of magnitude of the standard skull conductivity value of $8 \times 10^{-6} \text{ S mm}^{-1}$, as, e.g., shown in [6].

There are now different strategies to compute the $\hat{\sigma}_{\text{skull}}$ defined in equation (4.5). In [11], it is suggested to define a set $\{\sigma_{\text{skull}}^1, \dots, \sigma_{\text{skull}}^K\}$ of potential skull conductivity values, evaluate the residual variance to the EEG data for each member of this set, and take the member with the minimal value as the optimized conductivity value. In [6], it is suggested to start with an initial set $\{\sigma_{\text{skull}}^1, \dots, \sigma_{\text{skull}}^K\}$ like in [11] and compute the minimizing index in this set, say i_0 . [6] then suggests to define a new set $\{\tilde{\sigma}_{\text{skull}}^1, \dots, \tilde{\sigma}_{\text{skull}}^K\}$ by sampling a small interval around $\sigma_{\text{skull}}^{i_0}$, and choose the conductivity from this set that minimizes the residual variance. Speaking more plainly, [6] thus suggests a multi-resolution scan with two resolution levels. [185], on the other hand, suggests to assume that the mapping $\sigma_{\text{skull}} \mapsto \text{rv}_{\text{Id}}(d^{\text{EEG}}, L(\sigma_{\text{skull}}))$ is unimodal and apply Brent's method to compute its unique global minimizer.

We deviate from [6] and [185], and suggest using the method as [11] does. The central motivation for modifying [11] is the time demand of the algorithm. To achieve a reasonable computation time, [11] checked only 11 different conductivities, which are relatively coarsely spread

over the range of possible conductivity values. In this way, accuracy of the optimization is traded for increased computational efficiency. By only doing a finer sampling on a smaller subset, [6] tries to achieve a higher accuracy without too much additional effort, while the iterative optimization in [185] aimed at reducing the number of positions that need to be sampled. We instead argue that, instead of trying to change the algorithm from [11], it is more fruitful to optimize the computational efficiency of its implementation. To the best of our knowledge, the algorithm was up until now only implemented sequentially, see, e.g., [54], and the file “Leadfield_generation_calibration/D_calibrate_skull_cond.m” therein. But evaluating $\sigma_{\text{skull}} \mapsto \text{rvId}(d^{\text{EEG}}, L(\sigma_{\text{skull}}))$ at one position does not depend on the evaluation at another position, and there is thus no reason not to parallelize this evaluation. We can thus, without any real difficulty, divide the computational effort by the number of CPU cores by simply wrapping the implementation of [11] in a *parallel for loop*. We can, in fact, even further optimize this by adapting the C++ implementation in DUNEuro in such a way that it supports running multiple conductivity distributions in parallel, which, under the hood, all share the same grid data structure, thus keeping the memory overhead minimal. This is exactly what we did, and the corresponding implementation is available on the DUNEuro gitlab server in the branch “pipeline_version”.²⁰ Utilizing this strategy on an AMD Ryzen Threadripper 3960X CPU, we could sample 96 conductivity values with 48 parallel threads in ca. 15 min, with a peak memory consumption of ca. 16 GB. Compare this, e.g., to the computation of an MEG transfer matrix, which for the same head model on the same CPU took ca. 30 min, with a peak memory consumption of ca. 19 GB. The result of the conductivity sampling is visualized in figure 4.8, and yielded an optimized skull conductivity of ca. $0.0000187 \text{ S mm}^{-1}$.

We have thus finally constructed a calibrated, anisotropic, personalized tetrahedral head model for P1106. With this model in hand, we then used the previously constructed 1.5 mm source space and computed EEG and MEG lead fields by employing the PEM local subtraction approach as described in definition 2.5.12, utilizing a transfer matrix approach as described in section 2.9.4 for an efficient computation. Additionally, a tDCS forward solution for all electrodes supported by our stimulation cap was computed based on the PEM tDCS forward problem as given in definition 2.4.1.²¹

4.2.2 Estimating the generators of interictal epileptic activity

P1106 underwent a total of 2500 s of passive combined EEG/MEG measurement at the IBB. The generated data was then shared with the patient’s attending epileptologist for marking of epileptic spike activity. This resulted in the marking of 680 epileptic spikes. These spikes were categorized into two different patterns, called *pattern 1* and *pattern 2*, by the epileptologist. With the spike markings in hand, we then filtered the MEG and EEG signals using the same preprocessing parameters as [10], i.e., we applied a highpass filter of 1 Hz, a lowpass filter of 100 Hz, and a notch filter at 50 Hz for the power line noise. The preprocessing was again performed in FieldTrip [146]. Following [108] and [5], spike marker positions were moved to local peaks of relevant channels, with the hope of compensating for inaccuracies in the placement of the spike markers. Following this, noisy markings were identified and removed from further processing by

²⁰This branch is available at https://gitlab.dune-project.org/duneuro/duneuro/-/tree/pipeline_version?ref_type=heads. In this branch, the implementation can be found in the file “duneuro/driver/fit-volume_conductor.hh”, in the function called “calibrationScan”.

²¹Given the content of section 2.9.4, one might wonder why we needed an additional computation for the tDCS forward solution after already having employed a transfer matrix approach for the EEG forward problem. The reason for this is that the tDCS cap allows for the use of more electrodes than the cap used for the EEG measurement.

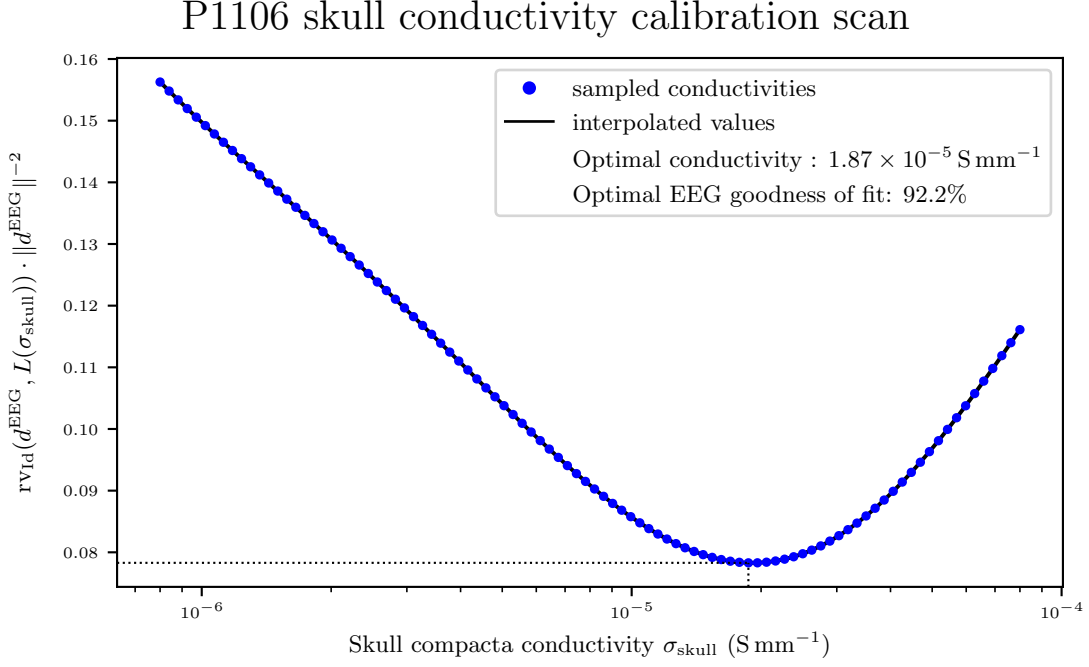


Figure 4.8: The so-called *calibration curve* for the skull conductivity of P1106. Using the dipole position reconstructed in figure 4.7, a set of candidate skull conductivity values is sampled and, for each such conductivity, the EEG goodness of fit of a dipole at the MEG position is computed. The skull conductivity is then estimated as the value that maximizes this goodness of fit, i.e., which minimizes the corresponding residual variance. Note that the default value suggested by [158] is $8 \times 10^{-6} \text{ S mm}^{-1}$, i.e., the optimized skull conductivity value is more than double the default skull conductivity.

visual inspection. Finally, all remaining spikes in pattern 1 and pattern 2 were averaged. The average of the pattern 1 spikes is visualized in figure 4.9, and the average of the pattern 2 spikes is visualized in figure 4.10.

In the EEG average of both pattern 1 and pattern 2, we see a topography that indicates a deep and radial dipolar activity. The MEG, on the other hand, does not show a dipolar topography, but instead shows an approximately quadrupolar topography in both pattern 1 and pattern 2. Furthermore, note that, in the MEG data, one can see in both patterns a strong rhythmic oscillation that is not visible in the EEG data. If one plots the MEG topography at the peaks of this oscillation, one sees a strong signal strength at the frontal MEG sensors. This signal is, however, not due to epileptic activity, but is instead due to a vibrating magnetized screw that was present inside the MEG chamber during the measurement, and that was only noticed at a later timepoint. The strong signal peaks in the MEG that are visible away from the 0 timepoint are thus of no relevance for the estimation of the underlying neural sources.

An area of the brain that generates interictal epileptic spikes is typically called an *irritative zone* [167]. The epileptologist who performed the markings on which figures 4.9 and 4.10 are based suggested that, for P1106, there are two irritative zones, one generating pattern 1 spikes, and one generating pattern 2 spikes. Furthermore, the epileptologist proposed that the pattern 1 spikes

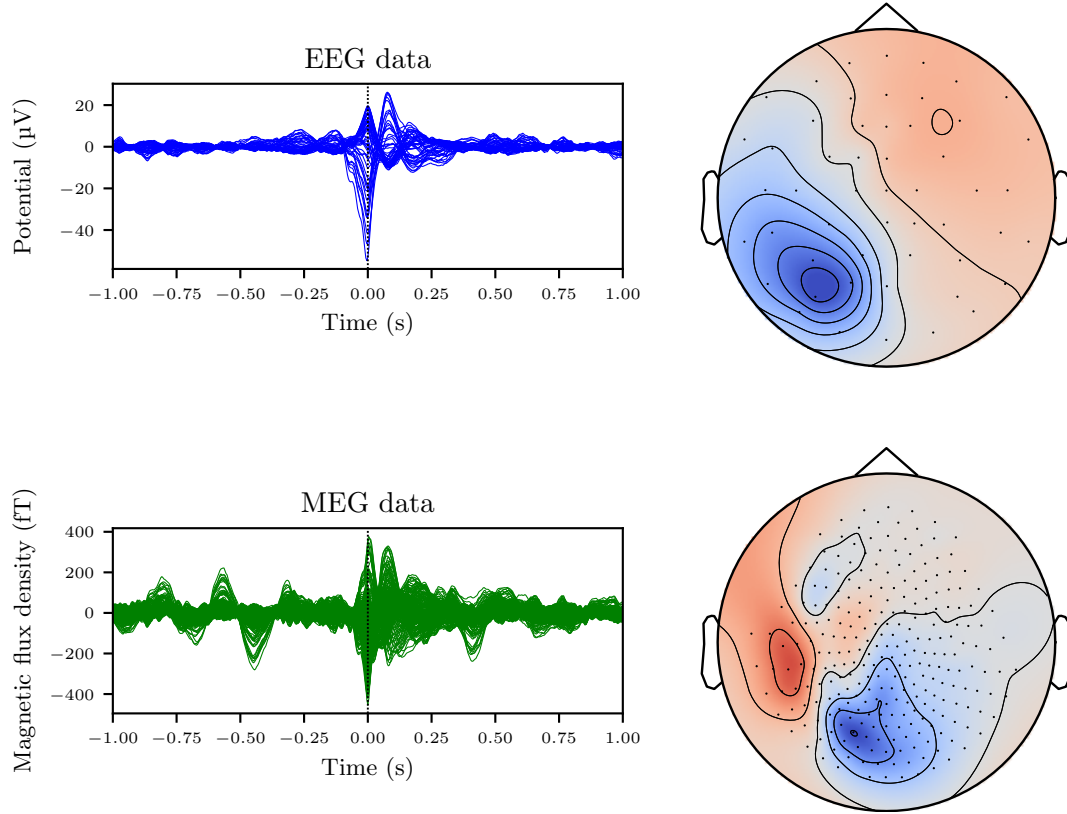


Figure 4.9: Visualization of the average of spikes classified as “pattern 1” by the epileptologist who performed the marking. This average was computed from 326 individual spikes. The topography was visualized in the same way as in figure 4.6.

are due to propagation of pattern 2 spikes. As the main goal of source estimation in epilepsy is the localization of the epileptogenic zone, one should try to minimize the distortions caused by the propagation of neural activity [115]. We thus primarily focused on pattern 2 spikes for the following investigation.

To this end, we tried to reconstruct the generators of the pattern 2 topographies by using a standard dipole scan as defined in 3.1.2, just as in the reconstruction of the somatosensory activity presented in the preceding subsection. For the EEG data, we were able to achieve a goodness of fit of ca. 97.6%, which is visualized in figure 4.11.

Furthermore, the corresponding location of the best-fitting dipole is concordant with a cortical malformation of P1106. In figure 4.12 we visualized all positions in the source space with an EEG goodness of fit of at least 95%, and when comparing these positions with figure 4.2, we see that the EEG reconstruction points to the patient’s polymicrogyria as the origin of the interictal epileptic spikes classified as pattern 2.

Trying to reconstruct the pattern 2 spike activity using the MEG data, on the other hand, fails miserably. For the sake of completeness, we have included the MEG dipole reconstruction for pattern 2 in figure 4.13. There, we see that the best-fitting MEG dipoles are those that

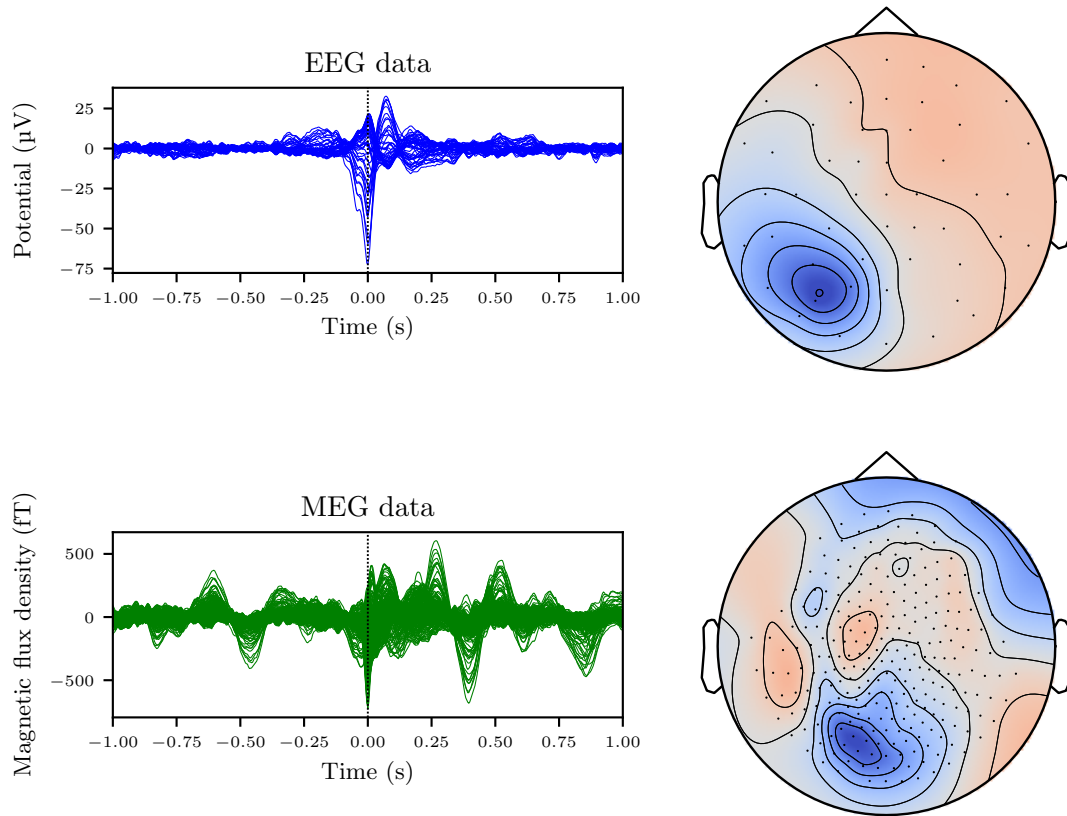


Figure 4.10: Visualization of the average of spikes classified as “pattern 2” by the epileptologist who performed the marking. This average was computed from 108 individual spikes. The topography was visualized in the same way as in figure 4.6.

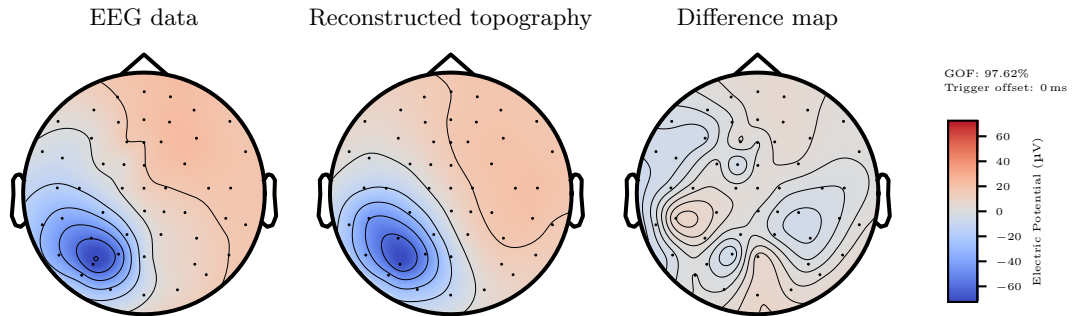


Figure 4.11: Visualization of the average EEG topography of the pattern 2 spike at its peak, the forward simulated EEG topography of the best-fitting dipole, and the difference between the two.

are positioned as inferior as possible, so that the positive MEG pole is outside of the sensor coverage and the MEG accordingly measures only the negative pole. In this way, at least the large negative pole in the MEG data is reconstructed, which gives the best achievable goodness

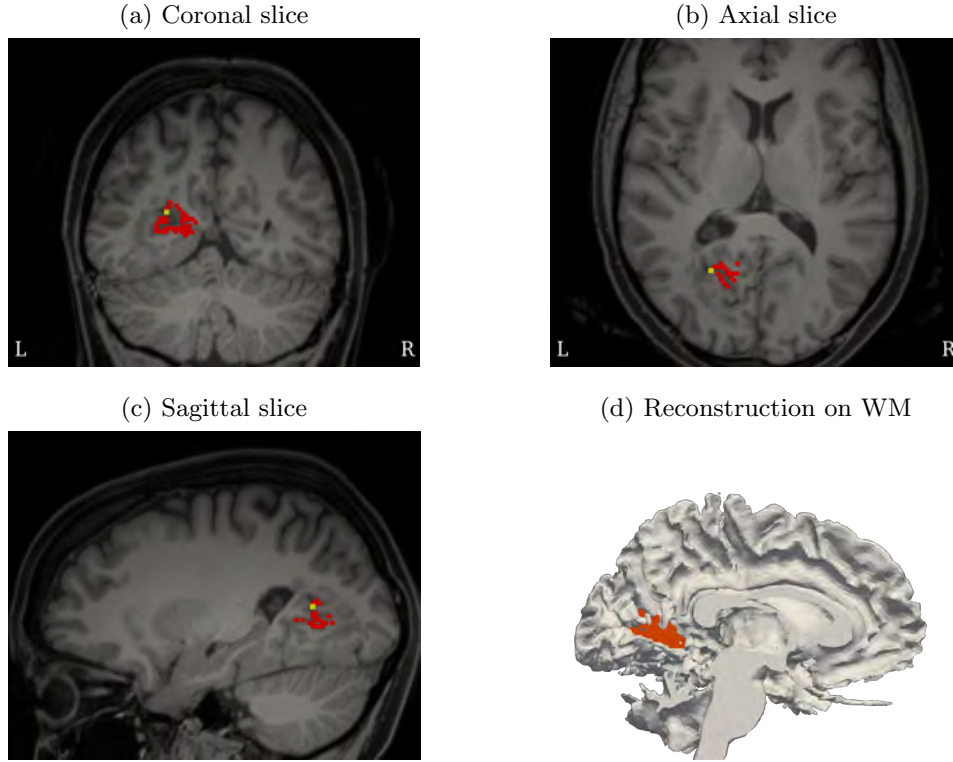


Figure 4.12: EEG dipole scan result for the average pattern 2 spike shown in figure 4.10. Each red box indicates a dipole position with a goodness of fit of at least 95%, while the yellow box indicates the position with the maximal goodness of fit of ca. 97.6%. In subfigures (a), (b), and (c), we respectively show all such dipole positions on a slice through the best-fitting dipole. In subfigure (d), we plot all high fitting dipole positions, i.e., a 3d point cloud instead of a 2d slice as in the other figures. Additionally, subfigure (d) visualizes all left hemisphere mesh elements assigned to the white matter compartment.

of fit of ca 78.8%. But when taking into account the EEG data, the anatomical information from the MRI, the seizure semiology, and the high residual variance in the MEG fit, a localization of spike activity in the cerebellum seems quite unlikely, which is our argument for saying that the MEG reconstruction failed. Furthermore, we want to emphasize that including the radial lead field component into the MEG reconstruction [56], trying a reconstruction at an earlier timepoint such as half of the rising flank [115, 10], or using the multipolar Venant approach [218] instead of the local subtraction approach for the computation of the MEG lead field all did not change this conclusion. Note, however, that, based on the literature, one would also expect that the MEG is not the optimal modality for this particular case. In [153], it is, e.g., shown that, when compared to the EEG, the MEG is insensitive to deep and radial sources. In figure 4.10, we see that the EEG is dominated by a deep and radial source, which is supported by the EEG source estimation shown in figure 4.12. With this in mind, we already a priori anticipate the EEG to produce a more dependable source estimation for the pattern 2 spikes than the MEG. Thus, based on the preceding discussion, we decided to disregard the MEG data for the source estimation of the interictal epileptic activity of P1106.

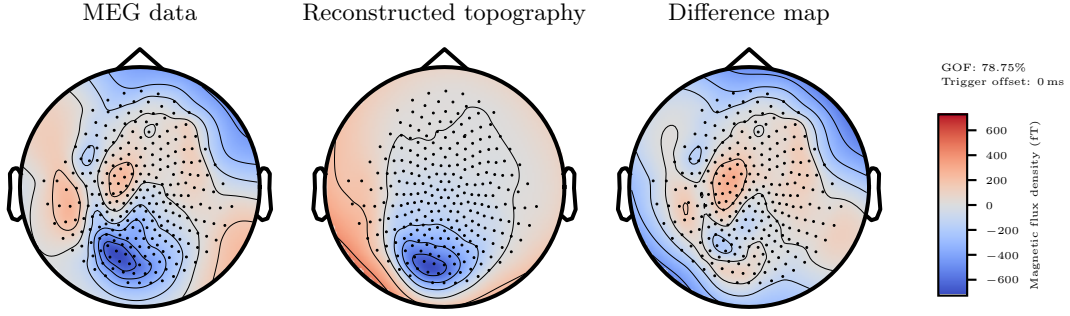


Figure 4.13: Visualization of the average MEG topography of the pattern 2 spike at its peak, the forward simulated MEG topography of the best-fitting dipole, and the difference between the two.

We also tried to estimate the generators of the pattern 1 spikes shown in figure 4.9, again using a dipole scan approach. In the EEG case, the estimation produced a result that was very similar to the one shown in figure 4.12, again pointing to the polymicrogyria as a likely candidate for the irritative zone. Similar to the case of pattern 2, the MEG estimation again peaked at a different location, and, following the same argument as for pattern 2, we again believe the EEG estimation to be more reliable.

In summary, our source estimation thus pointed at the polymicrogyria, shown in figure 4.2 (a), as a plausible candidate for the irritative zone. This evaluation was shared with the patient's epileptologist, who agreed that the polymicrogyria is a reasonable estimate for the patient's irritative zone. In agreement with the epileptologists, it was further decided to use the dipole with the highest goodness of fit from the EEG dipole scan of pattern 2 to define a target for the tDCS montage optimization.

4.3 TDCS stimulation of an epilepsy patient

4.3.1 Description of the stimulation montages

Before performing a tDCS stimulation, one first needs to define a stimulation montage. We discussed approaches for computing such montages in section 4.1. Based on our investigations in that section, we decided to utilize equation (4.2), i.e., the *modified D-CMI approach*, for the generation of an optimized current injection pattern for P1106.

If we want to employ equation (4.2), we need to define a target orientation $o \in \mathbb{R}^3$. Following [5], we aimed to *inhibit* the generation of epileptic spikes by trying to inject the current in such a way that it is *anti-parallel* to the estimated orientation of the interictal epileptic activity. If now $M \cdot \delta_{x_0}$ denotes the dipole with the highest EEG goodness of fit for the average pattern 2 spike from figure 4.10, we thus chose $o := -\frac{M}{\|M\|_2}$ for the definition of the modified D-CMI functional.

Based on our ethical clearance, we defined $C_{\text{total}} := 4 \text{ mA}$ and $C_{\text{max}} := 2 \text{ mA}$ for the constraint parameters in equation (4.2). For the determination of λ in (4.2), we followed the strategy suggested by [106] and [54]. In this strategy, we start at $\lambda = 0$ and iteratively increase λ in small steps until the point where for the first time the stimulation montage $\hat{I}_{\text{stim}}^\lambda$, as it is defined in (4.2), contains at least 9 electrodes over the activity threshold. The stimulation montage corresponding

Table 4.1: tDCS stimulation montages for P1106. In the leftmost column, electrodes are identified by their label in the 10-10 system for EEG. The second column contains the montage we originally intended to apply for the stimulation, while the third column contains the stimulation that was actually applied. The fourth column contains the stimulation montage that was applied during the sham stimulation. The final row lists the directionality of the individual stimulation montages. Furthermore, remember that a positive current is interpreted as current flowing into the head, while a negative current is interpreted as current flowing out of the head.

Montage	Planned	Actually applied	Sham
PO7 injection (mA)	-2.00	-1.00	-2.00
P7 injection (mA)	-2.00	-1.00	1.37
C4 injection (mA)	0.09	0.10	-2.00
T8 injection (mA)	0.84	0.44	0.04
FC6 injection (mA)	1.17	0.47	0.84
CP6 injection (mA)	0.04	0.10	0.49
FC4 injection (mA)	0.49	0.29	0.09
C6 injection (mA)	1.37	0.60	1.17
Directionality at target (mA m^{-2})	181.0	90.48	30.38

to the last λ -value for which at most 8 electrodes are above the activity threshold was then chosen as the basis for the tDCS stimulation of P1106. Following [54], we chose an activity threshold of 0.01 mA, which resulted in selecting $\lambda = 0.21$ for the final tDCS montage.

In table 4.1, we have described the stimulation montage we intended to apply, the stimulation montage we actually ended up applying (as described in the following paragraph), and the stimulation montage used in the sham stimulation. Additionally, we have included the directionality, i.e., the expression $\langle j_I(x_0), o \rangle$ as it is defined in section 4.1, for each stimulation montage.

Here, the stimulation montage we wanted to apply is the result of the modified D-CMI optimization (4.2) with the parameters defined above. This stimulation, however, exceeded the pain tolerance of P1106. In an attempt to reduce the pain, lidocaine was applied to the patient's skin below the stimulation electrodes to locally anesthetize the skin and thus reduce the pain caused by the stimulation. But even after applying the local anesthetic, the originally planned stimulation montage still exceeded the patient's pain threshold. At this point, the only remaining option was to reduce the amperage of the stimulation. As the patient had already terminated multiple stimulation attempts due to pain on the same day, we wanted to reduce the amperage by a sizable enough amount to achieve a bearable stimulation. We thus decided to halve the applied current at each electrode. This resulted in a stimulation that the patient deemed tolerable. However, in addition to the problem of finding a bearable stimulation, we also encountered technical difficulties. Namely, the stimulator would occasionally randomly stop the stimulation. A consultation with colleagues using the same stimulation device led to the hint that using electrodes with an amperage of below 0.1 mA increases the likelihood of an early termination of the stimulation. As the stimulation was performed in Bochum, it was not a feasible option to drive back to Münster and redo the optimization with new side constraints. Hence, we divided all planned amperages in half and then by hand redistributed the currents such that each electrode applied at least 0.1 mA. For this manual redistribution, we ensured that the relative ordering of the stimulation

intensities between the individual electrodes stayed identical. This finally yielded a montage that was both tolerable for the patient and did not cause the stimulation device to fail early on in the stimulation. The corresponding amperages are listed in the column with the title “Actually applied” in table 4.1. We want to emphasize that even though the redistribution to reach 0.1 mA was done “by hand”, the directionality of the actually applied montage is only marginally smaller than half of the directionality of the planned montage. Thinking reciprocally via theorem 4.1.4, this reflects the fact that the positive pole of the average pattern 2 spike is spread out over a large region of the right hemisphere of the head, as can be seen in figure 4.11.²²

Following [5] and [55], we decided to apply a so-called *active sham stimulation*. The purpose of this sham stimulation is to investigate to what degree the effects of the actual stimulation might be due to a placebo effect by comparing the actual stimulation to this sham stimulation. This is achieved by designing the sham stimulation in such a way that it leads to a strongly reduced directionality in the target, while simultaneously generating skin sensations that are similar to the actual stimulation. Ideally, the patient then cannot tell if the actual stimulation or the sham stimulation is applied. If the reader looks at the sham montage in table 4.1, they will see that the applied amperages are a permutation of the planned stimulation montage. Thus, the sham stimulation leads to the same skin sensations as the planned montage, but distributed differently over the selected electrodes. If the planned stimulation had been applied as originally intended, the sham montage would have yielded a directionality reduction in the target of ca. 83.21%. Note that this is comparable to the previously applied PerEpi stimulation montages, e.g., [5], where the sham stimulation results in a directionality reduction of ca. 85.71%, and [55], where the sham stimulation reduced the directionality by ca. 89.12%. However, the stimulation could not be performed as planned, and the directionality of the actual stimulation, which was applied after the sham stimulation, was reduced by half, so that in the end the sham stimulation only yielded a ca. 66.42% reduction in directionality when compared to the actually applied stimulation.

We have now described the stimulation montages that were applied (or where an application was tried) in the stimulation of P1106. In figure 4.14, the reader can find a visual summary of the stimulation montages from table 4.1, which in particular includes a comparison against the target EEG lead field.

4.3.2 Results of the stimulation

We followed the study design as outlined in [5], section 2.8. Concretely, on each stimulation day, the patient first underwent an EEG measurement for ca. 60 min. Then, the patient received two tDCS stimulation blocks with a duration of 20 min each, with a 20 min pause in between. Following this, the patient again underwent an EEG measurement of ca. 60 min. In a deviation from [5], this was not done for a total of 10 days, but instead only for a total of 4 days. This change was introduced after it became clear from the previous PerEpi stimulations, i.e., [5] and [55], that the burden on our clinical partners became too high, both financially, as the stimulated patient is occupying a hospital bed in an epilepsy unit for the whole duration of the stimulation, as well as with regard to the time needed for the evaluation of the stimulation results, as this requires the participating epileptologists to mark multiple hours of EEG data, which they need to perform alongside their already high clinical workload. Thus, we reduced the duration of the stimulation to two days for the actual stimulation and two days for the sham stimulation.

²²If the same stimulation device is again used in the future, and additionally some form of D-CMI approach is applied, we would advise to use an activity threshold of 0.1 mA already during cap optimization, to minimize the risk of an early termination of the stimulation.

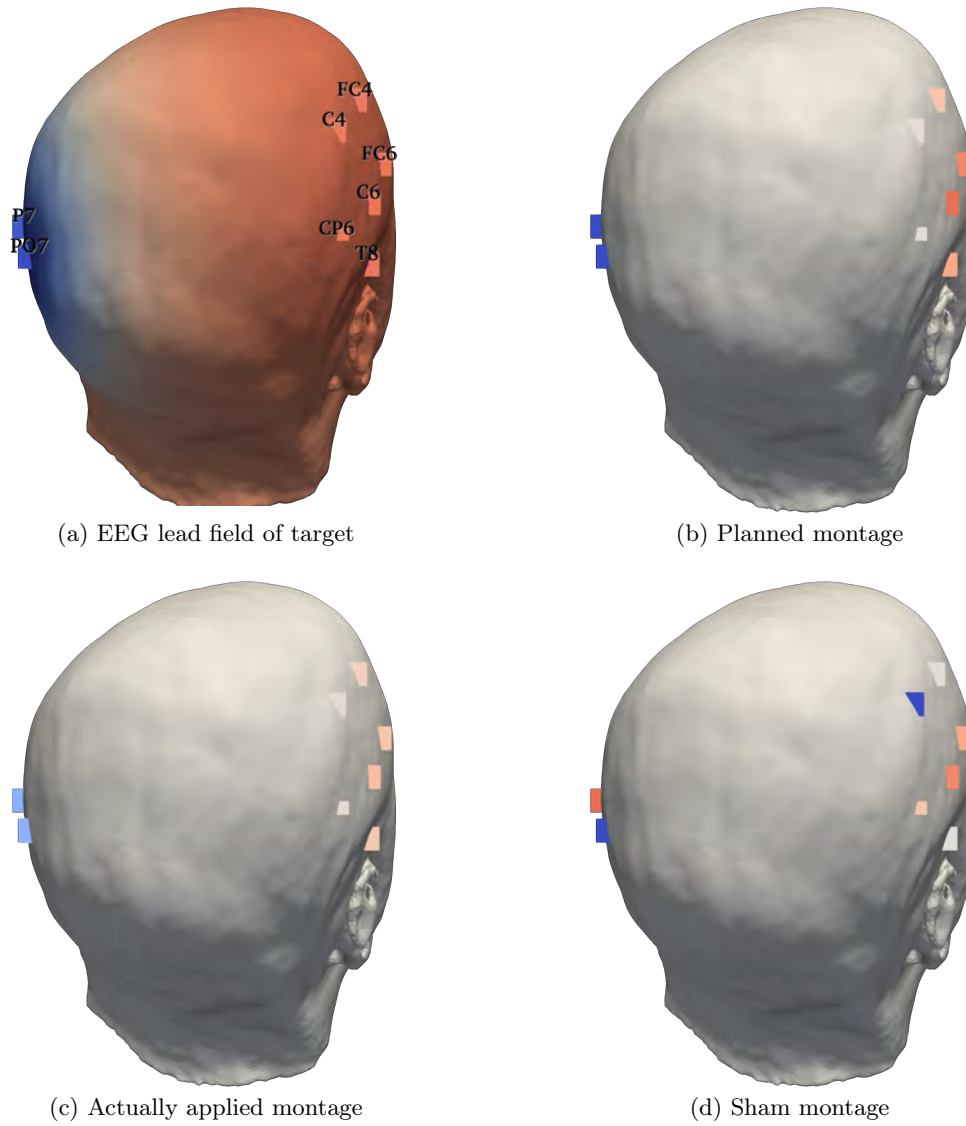


Figure 4.14: Visualization of the relevant stimulation montages for P1106, plotted on the FEM mesh. For the concrete numerical values of the applied currents, we refer to table 4.1. Figure (a) contains the EEG lead field of the inverted stimulation target, i.e., the lead field of the reconstructed pattern 2 spike from figure 4.11. By theorems 4.1.3 and 4.1.4, the montage maximizing the directionality in the target is the one that places as much amperage as possible on the poles of this lead field. By comparing a stimulation montage against this lead field, we can thus quickly check if the montage “makes sense”. Additionally, (a) also associates to each electrode position its label in the 10-10 system. Figures (b), (c), and (d) visualize the montages from table 4.1, and all share the same color mapping, which is most blue for -2 mA , and most red for 2 mA . Figure (a) uses a separate color mapping, which maps negative electric potential values to blue and positive electric potential values to red.

The first stimulation block was the sham stimulation and was performed in December 2024. The second stimulation block was the actual stimulation and was performed in February 2025. The large time difference between the stimulation blocks was chosen to ensure a proper wash-out of stimulation effects and thus prevent effects of the first stimulation block on the second [240]. Furthermore, we want to emphasize that neither the patient receiving the stimulation nor the nurses and medical doctors administering the stimulation were informed about whether a stimulation day consisted of a sham stimulation or an actual stimulation. The questionnaires filled out by the patient after the stimulation confirmed that the patient could not distinguish between the true stimulation and the sham stimulation. The study can thus be classified as double-blinded.

The experiment resulted in a total of 8 measured EEG blocks with a length of ca. 60 min each, as we measured two EEG blocks per day, for two days where the actual stimulation was applied and for two days where the sham stimulation was applied. To preserve the blinding, these blocks were then truncated to a uniform length, had their order randomly shuffled around, and were assigned random names. In this randomized form, the EEG blocks were then shared with three participating epileptologists, called epileptologist 1, 2, and 3 in the following. Each of these epileptologists then investigated the individual EEG blocks and annotated the data at timepoints where they identified epileptic spikes. These so-called *markings* were then reported back to us, where we reversed the anonymization and thus arrived at the result of the stimulation. In table 4.2, the reader can find the number of epileptic spikes marked by epileptologist 1, in table 4.3 the corresponding markings of epileptologist 2, and in table 4.4 the markings of epileptologist 3.

First of all, note that even though each of the three markers is a trained epileptologist actively involved in clinical duties in a university clinic, they arrive at strongly different spike counts. Epileptologist 2 marked, on average, ca. 53% fewer spikes than epileptologist 1, and epileptologist 3 marked, on average, ca. 25% fewer spikes than epileptologist 1. This is consistent with what one would expect based on the literature, as the concrete choice on what kind of brain activity is actually an epileptic spike, and what activity is noise or background brain activity, is, to some degree, subjective [234]. This is also consistent with our own experience in the previous PerEpi stimulation blocks in [5] and [55], where we also observed that different epileptologists can, for the same EEG dataset, arrive at quite different conclusions with regard to the number of spikes present in the data.

But while spike counts might vary strongly between epileptologists, they are far from being stochastically independent. This is illustrated in figure 4.15, where we show scatter plots comparing the number of spikes as marked by the different epileptologists. There, we see that for each pair of epileptologists, i.e., 1 and 2, 1 and 3, and 2 and 3, there is a very high correlation of more than 0.95 between the marked number of spikes. Thus, while the epileptologists might not agree on the concrete number of spikes that are present in a dataset, they generally agree on whether a given EEG block contains more or less epileptic activity than a reference EEG block.

And they agree that the tDCS stimulation did not decrease the amount of epileptic activity visible in the patient's EEG data. On the contrary, looking at the stimulation results in tables 4.2, 4.3, and 4.4, we see that each epileptologist in fact detected an increase in epileptic activity after the true stimulation. Furthermore, we also see that, on the second day of sham stimulation, the amount of epileptic activity also increased in each of the individual markings. If we measure the change in epileptic activity as the ratio of the number of spikes after the stimulation and the number of spikes before the stimulation, we see that the increase in spikes on the second sham

	True stimulation		Sham stimulation	
	Day 1	Day 2	Day 1	Day 2
Before	194	478	147	65
After	526	718	136	245

Table 4.2: Number of epileptic spikes marked by epileptologist 1 in each measured EEG block. On each stimulation day, an EEG block was measured before and after the tDCS stimulation. A total of four stimulation days were carried out, two for the sham stimulation and two for the true stimulation, resulting in a total of 8 measured EEG blocks.

	True stimulation		Sham stimulation	
	Day 1	Day 2	Day 1	Day 2
Before	128	292	66	21
After	195	374	60	86

Table 4.3: Number of epileptic spikes marked by epileptologist 2 in each measured EEG block. The markings were performed on the same EEG blocks that are underlying table 4.2.

	True stimulation		Sham stimulation	
	Day 1	Day 2	Day 1	Day 2
Before	189	421	66	31
After	460	569	128	140

Table 4.4: Number of epileptic spikes marked by epileptologist 3 in each measured EEG block. The markings were performed on the same EEG blocks that are underlying table 4.2.

stimulation day is even larger than the increase in spikes on the two stimulation days. Thus, there was not only no reduction in epileptic activity after the application of the stimulation, but the increase in epileptic activity that can be seen on the true stimulation days is in the same range as observed on the second sham stimulation day. Thus, the stimulation results in tables 4.2, 4.3, and 4.4 give no indication that the true stimulation had a different effect than the sham stimulation, i.e., we could not work out any evidence for the efficacy of tDCS stimulation in this experiment.

What we can, however, see is that the number of spikes after the stimulation increases, on average over all epileptologists and stimulation days, by ca. 125% when compared to the number of spikes before the stimulation. This suggests that the stimulation had an effect on the patient's epileptic activity, which, due to the comparison with the sham stimulation, cannot be expected to be due to an effect of the electric current on the brain. We instead suspect that the increase in epileptic activity can be attributed to the stress caused by the stimulation. It has been demonstrated

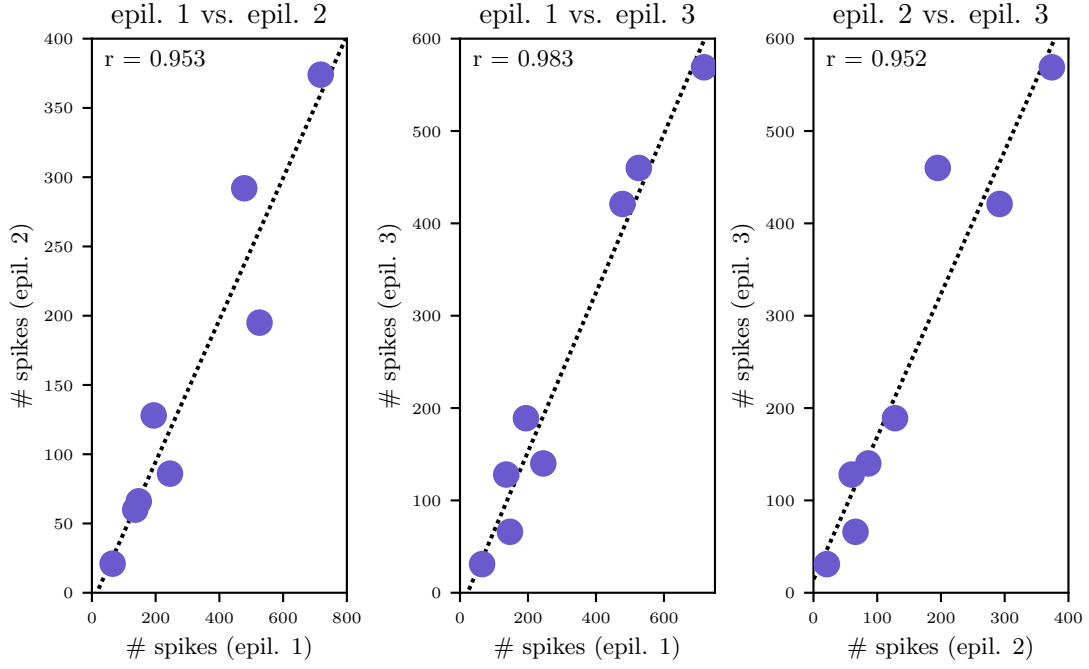


Figure 4.15: Comparison of the number of marked spikes between different epileptologists. In the leftmost plot, we compared epileptologists 1 and 2 by taking each corresponding pair of values from the tables 4.2 and 4.3 and plotting the resulting 8 pairs in a scatter plot as blue dots. Additionally, we computed the sample correlation coefficient, shown in the upper left corner of the plot, and the linear regression based on a least squares approach, shown as a dotted line. We refer to [168] for details on this computation. The same procedure was repeated for the comparison of epileptologists 1 and 3, yielding the middle plot, and for the comparison of epileptologists 2 and 3, yielding the rightmost plot.

that the stress level of an epilepsy patient is associated with both the seizure frequency and the spike frequency [207, 32]. We believe that the stimulation caused a significant amount of stress for the patient, which resulted in the increase in epileptic activity after the stimulation that can be seen in tables 4.2, 4.3, and 4.4. In section 4.3.1, we described the difficulties that occurred during the stimulation blocks. The most notable of these was that the true stimulation had to be repeatedly stopped, as it exceeded the patient's pain tolerance. Only after multiple attempts of first applying an anesthetic and then reducing the amperage was the stimulation deemed tolerable. Additionally, note that even if a stimulation is "tolerable", it is still far from pleasant, and the patient had to endure 40 min of it. We assume that these circumstances had a negative effect on the stress levels of both the patient and the nurse administering the stimulation. Furthermore, in the sham stimulation, we encountered technical difficulties, where the stimulation device would seemingly randomly stop the stimulation. This certainly caused stress on our side and on the side of the nurses, and it is likely that this also had an effect on the patient's stress level.

4.4 Discussion and patient outlook

In this chapter, we described our pipeline for the construction of optimized stimulation montages for the tDCS stimulation of focal epilepsy patients, and the results of applying this stimulation to a patient. We started with a discussion and comparison of different montage optimization algorithms in section 4.1, finally deciding to use the approach defined by equation (4.2), i.e., a modification of the D-CMI algorithm from [106], for the final montage computation. After having discussed the methodological basis of the stimulation approach, we introduced the epilepsy patient who agreed to partake in our tDCS study. For this patient, we then utilized all modeling modalities available at the IBB to generate a personalized FEM head model, including a segmentation derived from T1- and T2-MRI images, DTI-derived anisotropic white matter conductivity tensors, a skull conductivity calibrated based on a somatosensory evoked response measured using combined EEG and MEG, and sensor positions registered to the individual head shape by means of a Polhemus-based digitization of the patient’s head surface. Additionally, the patient underwent a resting-state combined EEG and MEG measurement at the IBB for a duration of ca. 42 min, which was shared with the patient’s attending epileptologist for the marking of interictal epileptic activity. The identified epileptic activity was then further used to estimate its underlying neural generators, which, in discussion with the patient’s attending epileptologists, was deemed reasonable. With the goal of inhibiting the interictal epileptic activity, a montage was optimized to inject a current that runs anti-parallel to the estimated neural generators of this activity. This montage was then applied, unfortunately with half the planned amperage, in a double-blinded and sham-controlled study. In this experiment, the goal of reducing the epileptic activity could not be achieved. In contrast, the epileptic activity was increased after the stimulation. This increase, however, was present in both the sham stimulation as well as in the true stimulation, implying that the increase was not due to an effect of the stimulation on neural populations, but rather due to outside factors. As outlined in the preceding section, we in fact believe that the stress caused by the stimulation led to an increase that eclipses any potential effects of the stimulation on the patient’s brain.

When discussing these results within our workgroup at the IBB, it was suggested that, given that the target lies relatively deep inside the head, as can be seen in figure 4.2 (a), it was somewhat unreasonable to even expect the stimulation to have any effect. We disagree with this assessment. Even though the target was relatively deep, the planned stimulation montage still achieved a directionality of 181 mA m^{-2} , as can be seen in table 4.1. And while this is somewhat lower than the directionalities achieved in the previous PerEpi stimulations, i.e., 210 mA m^{-2} in [5] and 258 mA m^{-2} in [55], it is still of a somewhat comparable magnitude. In the wider literature, it has been shown that even much weaker fields can have a measurable effect on neural networks. In [65] it is, e.g., demonstrated that an electrical field with a peak amplitude of $295 \text{ } \mu\text{V mm}^{-1}$ has a detectable effect on neural networks. Using the default SimNIBS gray matter conductivity of 0.275 S m^{-1} , as given in [158], this translates to a current density of ca. 80 mA m^{-2} .²³ One might now respond that there is probably some gap between an electric field that has a detectable effect on neural networks in isolated brain slices under laboratory conditions and an electric field that has a measurable effect on live patients under clinical conditions. The 80 mA m^{-2} can thus be thought of as more of a “lower bound” that the current density in the target should surpass. Furthermore, in [105] it was demonstrated that in a D-CMI stimulation study with 13 participants a mean directionality of 126 mA m^{-2} , with a standard deviation of 70 mA m^{-2} , was sufficient to

²³This is essentially the same line of reasoning as it is employed in [183], where the authors, however, due to a slightly different interpretation of [65] and a different choice of default gray matter conductivity, arrive at a value of ca. 60 mA m^{-2} . We thus want to emphasize that our value of ca. 80 mA m^{-2} , or any other similarly derived value, is more of a “back of the envelope” calculation, and should be taken with a grain of salt.

achieve a significant increase in source amplitude in the stimulated brain areas. Additionally, it should be emphasized that only looking at the directionality generated by a montage completely ignores the component of the induced current density flowing orthogonally to the target. Based on physiological reasoning, as it is, e.g., described in [121], one can, however, also expect that these orthogonal currents will have some degree of influence on the effect of the stimulation. Given this context, we do not believe that it was a priori unreasonable to investigate whether the planned stimulation montage from table 4.1 would have an effect on the patient’s epileptic activity.

However, as described in the preceding section, we in the end only applied half the amperage of the planned stimulation montage, due to the planned stimulation montage exceeding the pain tolerance of the patient on the day of the stimulation. This resulted in the application of a montage with a directionality of ca. 90 mA m^{-2} , which is only slightly above the “lower bound” of 80 mA m^{-2} mentioned in the last paragraph. Our actual stimulation strength was thus barely in the range of what is needed to detect effects under ideal conditions, and it is thus not particularly surprising that we were unable to detect any effects that are different from what is visible in the sham condition. The study presented in this chapter thus constitutes the first case where a PerEpi stimulation did not achieve a reduction of epileptic activity, as the first stimulation in [5] resulted in a significant reduction of the number of epileptic spikes, and the second stimulation in [55] resulted in a significant reduction of the number of seizures.

Another point that is worth discussing is the value of the skull conductivity that was used in the EEG/MEG/tDCS simulations. Remember that the skull conductivity was calibrated following the strategy from [11], with the resulting calibration curve shown in figure 4.8. As we cannot immediately measure the skull conductivity, this procedure implicitly estimates the skull conductivity by trying to select a value such that the MEG and EEG dipole reconstructions of a somatosensory evoked response match. Following [11] and [92], the reasoning for this is that the MEG is much less affected by conductivity uncertainties than the EEG, and one can thus, even in the presence of large conductivity uncertainties, reliably perform a source estimation based on the MEG data. As the EEG measurements and the MEG measurements are based on the same underlying neural activity, their reconstructions should also match, as long as the activity is visible in both modalities. We, however, cannot expect that this approach will precisely estimate the skull conductivity for a number of different reasons. First of all, there is noise in the measurements, there will be segmentation errors, the numerical simulations will yield some inaccuracies, and the skull will not have a constant conductivity due to, e.g., an inhomogeneous density. But more importantly, the skull is, of course, not the only tissue whose exact conductivity we do not know, and solely fitting the skull conductivity while keeping the remaining tissue conductivities fixed to default values will result in the skull conductivity value compensating for the inaccuracies in the other tissue conductivities, as was demonstrated in [185]. Thus, the “calibrated” skull conductivity should not be viewed exclusively as an estimation of the skull conductivity, but more as a relatively coarse single-parameter correction for all the errors discussed above. With this in mind, we see that even after skull conductivity calibration, there is still quite a lot of uncertainty in the actual skull conductivity value, and also in the remaining tissue conductivities, such as the scalp conductivity. This conductivity uncertainty, however, directly translates into an uncertainty with regard to the current density magnitude in the tDCS target position, as has been shown in [179] and [183]. The current density magnitude at the tDCS target is, on the other hand, directly related to the directionality of the stimulation montage. In summary, this yields that the directionality values that are reported in table 4.1, and by the same argument also those reported, e.g., in [5], [55], and [105], come attached with a certain degree of uncertainty, and one should be wary of not overinterpreting the concrete values

of these metrics.

We also want to emphasize that the experiment presented in this chapter utilized our newly developed local subtraction approach for the EEG and MEG forward problems, as it is described in 2.5.12 and 2.5.17. We can, however, not claim that using the local subtraction approach had a significant impact on the source estimation. If we instead use a lead field computed using the multipolar Venant approach from [218], which we would consider the current state-of-the-art for MEEG FEM forward modeling, and apply this lead field in a dipole scan reconstruction of the pattern 2 spikes from figure 4.10, we see that the corresponding goodness of fit is maximized in the same element of the source space as it is for the local subtraction lead field. Thus, the local subtraction approach and the multipolar Venant approach yield the exact same estimated source position. From a negative point of view, one might wonder if the considerable effort that has gone into the development of the local subtraction approach was worth it. From a more positive angle, we see that two different forward modeling approaches, based on quite different strategies, arrive at the same conclusion. The two approaches thus validate each other, and their concordance increases the confidence in the result of the source estimation. We hope that the question of whether there is some systematic advantage or disadvantage of the local subtraction approach in comparison with the multipolar Venant approach in the context of MEEG-based source estimation will be further investigated by future research.

As a final consideration, we want to discuss the impact of the experiment on the patient. First of all, it is important to note that the goal of the experiment was not to offer tDCS stimulation as a continuing treatment option to the patient, but rather to investigate the potential of individually targeted and optimized tDCS stimulation in focal epilepsy, on the one hand as a potential treatment option, but on the other hand also as an additional tool in the diagnosis of epilepsy.

One central problem in the treatment process of focal epilepsy patients is the localization of the epileptogenic zone, i.e., the part of the brain that needs to be removed to achieve seizure freedom [167]. This “localization”, however, is, at the time of writing this thesis, not a problem with an exact solution, but rather consists of gathering (hopefully) converging evidence from multiple sources, such as seizure semiology or MEEG data, and building up enough confidence in the localization estimate such that the risk of a surgical procedure seems justified. Furthermore, as brain tissue can only be resected exactly once, surgeons generally try to remove as little brain tissue as possible to minimize post-operative neurological deficits [230]. It is thus of great interest to estimate the epileptogenic zone in such a way that the estimate is both as small as possible, but at the same time also has a high chance of being related to the generation of epileptic seizures. One possible contribution of a tDCS stimulation now consists of being a minimally invasive strategy to get information on the potential effect of removing certain parts of the cortex. If a tDCS stimulation that is individually optimized to target a certain brain region achieves a significant reduction in epileptic activity, this is a strong indicator that the targeted brain region is indeed part of the epileptogenic zone and that a resection is likely to have an effect on the patient’s epileptic activity. Especially in cases with low confidence in the estimated epileptogenic zone, or in cases where there are multiple candidates for the epileptogenic zone, a tDCS stimulation might be a worthwhile additional diagnostic component in the pre-surgical planning of a resection.

Another possible use of tDCS lies in the assessment of whether a patient is a suitable candidate for the implantation of a brain stimulation device. While it has been repeatedly demonstrated that tDCS stimulation can achieve a significant reduction of epileptic activity, the corresponding studies also report a high rate of non-responders and a large inter-subject variability, see, e.g.,

[104, 188]. This high variability is in fact a well-known phenomenon in the brain stimulation community [118, 122], and the expectation that at least part of this variability is due to individual differences in head anatomy, as it is, e.g., articulated in [112], was one of the main motivations for the PerEpi project, where the hope was that personalized stimulation montages might result in more reliable tDCS stimulation outcomes. Now, if a patient is considering the surgical implantation of a brain stimulator, it would be highly desirable to know beforehand whether the patient would even respond to the stimulation, as this might, on the one hand, prevent the patient from undergoing the burden of a futile surgery, and, on the other hand, avoid potential side effects of a long-term electrical brain stimulation on the patient. A tDCS stimulation prior to surgery can contribute to estimating the effect of brain stimulation on the patient, which can then inform the choice of potentially proceeding to more invasive measures. Thinking further, the outcome of a tDCS experiment might not only influence whether the implantation of a stimulator is recommended, but can also affect where the stimulator is placed during surgery.

The discussion of this last paragraph, in fact, applies to the patient we have presented in this chapter. At the point of writing this thesis, the patient is considering the implantation of an EASEE device, as it is used in the studies [187] and [188]. One goal of the stimulation experiment was thus in particular to examine if the patient might profit from the implantation of an EASEE stimulator,²⁴ which consists of 5 electrodes placed directly on the patient’s skull in a pseudo-Laplacian arrangement, with a more detailed description of the stimulator geometry and the stimulation paradigm available in [188]. To assess how an EASEE stimulation would relate to the montages presented in table 4.1, we performed a numerical simulation in a multilayer-sphere model emulating the EASEE setup, utilizing the complete electrode model as described in definition 2.4.2. This yielded the result that, due to the depth of the target in combination with the pseudo-Laplacian arrangement of the electrodes, we expect the current density generated by the EASEE stimulator to have only slightly larger magnitude than the current density corresponding to the “actually applied” montage from table 4.1.²⁵ Simultaneously, the current density magnitude generated by the EASEE stimulation setup in the brain regions directly below the skull is almost 20 times larger than the magnitude in the deep target position.²⁶ As the “actually applied” montage had no noticeable effect on the measured epileptic activity of the patient, and the large current density magnitude at the more superficial brain regions raises strong concerns with regard to potential side effects, we thus, in a discussion with the patient’s attending epileptologist, arrived at the conclusion that we do not believe the patient to be a suitable candidate for the implementation of an EASEE stimulation device. At the point of writing this thesis, the final decision of the patient on whether to proceed with the implementation is still open.

4.5 Conclusion

Transcranial direct current stimulation is a promising neuromodulation technique, which is, however, plagued by a substantial interindividual variability in efficacy. While previous stimulations in our PerEpi project showed the potential of reducing this variability in focal epilepsy stimulation by means of personalized stimulation montages based on individualized realistic head

²⁴<https://easee.precisis.de/de/>

²⁵The “actually applied” montage yielded a current density in the target with a magnitude of ca. 124 mA m^{-2} , while an EASEE stimulation with an injection of strength 4 mA at the central electrode and an extraction of 1 mA at the surrounding electrodes yielded a current density with a magnitude of ca. 128 mA m^{-2} at a target position with a distance of 28.5 mm from the skull boundary placed directly below the central EASEE electrode in a standard four layer sphere model.

²⁶In the same 4 layer sphere model emulating the EASEE setup that was used in the preceding footnote, the current density at the position directly below the central EASEE electrode with a distance of 3 mm from the skull boundary has a magnitude of ca. 2450 mA m^{-2} .

models coupled with state-of-the-art FEM simulations, we saw in this thesis that there are also cases where, even after employing our most modern approaches in forward modeling, source estimation, and cap optimization, we can still fail in reducing the patient's epileptic activity. Individual factors such as stress might eclipse any stimulation effect, not completely controllable factors during the stimulation might result in a suboptimal execution of the experiment, or the anatomical structure of the patient might prevent the application of a sufficient dosage due to the risk of side effects. Thus, while we believe in the benefit of personalization in tDCS stimulation, it is also clear that individually optimized stimulation caps will not yield a "miracle cure", but will instead "only" cut down the tDCS outcome variability to some degree. Finally, it should also be emphasized that even a tDCS stimulation without a noticeable effect can have a tangible positive impact on a patient, as the information of whether a patient responds to brain stimulation can influence the further treatment process, e.g., in the decision on whether to proceed with the surgical implantation of a brain stimulation device.

Bibliography

- [1] J. P. AHRENS, B. GEVECI, AND C. C. LAW, *ParaView: An end-user tool for large-data visualization*, in The Visualization Handbook, 2005.
- [2] M. AINSWORTH AND J. COYLE, *Hierarchic finite element bases on unstructured tetrahedral meshes*, International Journal for Numerical Methods in Engineering, 58 (2003), pp. 2103 – 2130.
- [3] P. ALLIEZ, C. JAMIN, L. RINEAU, S. TAYEB, J. TOURNOIS, AND M. YVINEC, *3D mesh generation*, in CGAL User and Reference Manual, CGAL Editorial Board, 6.0.1 ed., 2024.
- [4] A. ANTAL, I. ALEKSEICHUK, M. BIKSON, J. BROCKMÖLLER, A. BRUNONI, R. CHEN, L. COHEN, G. DOWTHWAITE, J. ELLRICH, A. FLÖEL, F. FREGNI, M. GEORGE, R. HAMILTON, J. HAUEISEN, C. HERRMANN, F. HUMMEL, J. LEFAUCHEUR, D. LIEBE-TANZ, C. LOO, C. MCCAIG, C. MINIUSSI, P. MIRANDA, V. MOLIADZE, M. NITSCHKE, R. NOWAK, F. PADBERG, A. PASCUAL-LEONE, W. POPPENDIECK, A. PRIORI, S. ROSSI, P. ROSSINI, J. ROTHWELL, M. RUEGER, G. RUFFINI, K. SCHELLHORN, H. SIEBNER, Y. UGAWA, A. WEXLER, U. ZIEMANN, M. HALLETT, AND W. PAULUS, *Low intensity transcranial electric stimulation: Safety, ethical, legal regulatory and application guidelines*, Clinical Neurophysiology, 128 (2017), pp. 1774–1809.
- [5] M. ANTONAKAKIS, F. KAISER, S. RAMPP, S. KOVAC, H. WIENDL, W. STUMMER, J. GROSS, C. KELLINGHAUS, M. KHALEGHI-GHADIRI, G. MÖDDEL, AND C. H. WOLTERS, *Targeted and optimized multi-channel transcranial direct current stimulation for focal epilepsy: An N-of-1 trial*, Brain Stimulation, 17 (2024), pp. 221–223.
- [6] M. ANTONAKAKIS, S. SCHRADER, ÜMIT AYDIN, A. KHAN, J. GROSS, M. ZERVAKIS, S. RAMPP, AND C. H. WOLTERS, *Inter-Subject Variability of Skull Conductivity and Thickness in Calibrated Realistic Head Models*, NeuroImage, 223 (2020), p. 117353.
- [7] F. ASHRAFZADEH, J. AKHONDIAN, N. HASHEMI, M. ESMAEILZADEH, A. GHANAEE, H. YAVARZADEH, S. IMANNEZHAD, N. SAEEDI ZAND, H. S. MIRZADEH, AND M. BEIRAGHI TOOSI, *Therapeutic impacts of transcranial direct current stimulation on drug-resistant epilepsy in pediatric patients: A double-blind parallel-group randomized clinical trial*, Epilepsy Research, 190 (2023), p. 107074.
- [8] ———, *Therapeutic impacts of transcranial direct current stimulation on drug-resistant epilepsy in pediatric patients: A double-blind parallel-group randomized clinical trial*, Epilepsy Research, 190 (2023), p. 107074.

- [9] K. AWADA, D. JACKSON, J. WILLIAMS, D. WILTON, S. BAUMANN, AND A. PAPANICOLAOU, *Computational aspects of finite element modeling in EEG source localization*, IEEE Transactions on Biomedical Engineering, 44 (1997), pp. 736–752.
- [10] U. AYDIN, J. VORWERK, M. DÜMPELMANN, P. KÜPPER, H. KUGEL, M. HEERS, J. WELLMER, C. KELLINGHAUS, J. HAUEISEN, S. RAMPP, H. STEFAN, AND C. H. WOLTERS, *Combined EEG/MEG Can Outperform Single Modality EEG or MEG Source Reconstruction in Presurgical Epilepsy Diagnosis*, PLOS ONE, 10 (2015), pp. 1–29.
- [11] U. AYDIN, J. VORWERK, P. KÜPPER, M. HEERS, H. KUGEL, A. GALKA, L. HAMID, J. WELLMER, C. KELLINGHAUS, S. RAMPP, AND C. H. WOLTERS, *Combining EEG and MEG for the reconstruction of epileptic activity using a calibrated realistic volume conductor model*, PLOS ONE, 9 (2014), pp. 1–17.
- [12] H. AZIZOLLAHI, M. DARBAS, M. M. DIALLO, A. E. BADIA, AND S. LOHRENGEL, *EEG in neonates: Forward modeling and sensitivity analysis with respect to variations of the conductivity*, Mathematical Biosciences and Engineering, 15 (2018), pp. 905–932.
- [13] P. BASTIAN, K. BIRKEN, K. JOHANNSEN, S. LANG, N. NEUSS, H. RENTZ-REICHERT, AND C. WIENERS, *UG—a flexible software toolbox for solving partial differential equations*, Comp. Vis. Sci, 1 (1997), pp. 27–40.
- [14] P. BASTIAN, M. BLATT, A. DEDNER, N.-A. DREIER, C. ENGWER, R. FRITZE, C. GRÄSER, C. GRÜNINGER, D. KEMPF, R. KLÖFKORN, M. OHLBERGER, AND O. SANDER, *The Dune framework: Basic concepts and recent developments*, Computers & Mathematics with Applications, 81 (2021), pp. 75–112. Development and Application of Open-source Software for Problems with Numerical PDEs.
- [15] M. BAUER, S. PURSIAINEN, J. VORWERK, H. KÖSTLER, AND C. WOLTERS, *Comparison study for whitney (raviart-thomas) type source models in finite element method based eeg forward modeling*, IEEE transactions on bio-medical engineering, 62 (2015).
- [16] E. BEGHI, *The Epidemiology of Epilepsy*, Neuroepidemiology, 54 (2019), pp. 185–191.
- [17] L. BELTRACHINI, *The subtraction reciprocity theorem for volume conductors*, To be submitted.
- [18] L. BELTRACHINI, *The analytical subtraction approach for solving the forward problem in EEG*, Journal of Neural Engineering, 16 (2019), p. 056029.
- [19] L. BELTRACHINI, *A Finite Element Solution of the Forward Problem in EEG for Multipolar Sources*, IEEE Transactions on Neural Systems and Rehabilitation Engineering, 27 (2019), pp. 368–377.
- [20] L. BELTRACHINI, *Sensitivity of the projected subtraction approach to mesh degeneracies and its impact on the forward problem in EEG*, IEEE Transactions on Biomedical Engineering, 66 (2019), pp. 273–282.
- [21] L. J. BINDMAN, O. C. J. LIPPOLD, AND J. W. T. REDFEARN, *Long-lasting Changes in the Level of the Electrical Activity of the Cerebral Cortex produced by Polarizing Currents*, Nature, 196 (1962), pp. 584–585.
- [22] G. BIRBECK, R. HAYS, X. CUI, AND B. VICKREY, *Seizure Reduction and Quality of Life Improvements in People with Epilepsy*, Epilepsia, 43 (2002), pp. 535–538.

- [23] M. BLATT, *A Parallel Algebraic Multigrid Method for Elliptic Problems with Highly Discontinuous Coefficients*, dissertation, 01 2010.
- [24] P. S. BOGGIO, S. P. RIGONATTI, R. B. RIBEIRO, M. L. MYCZKOWSKI, M. A. NITSCHKE, A. PASCUAL-LEONE, AND F. FREGNI, *A randomized, double-blind clinical trial on the efficacy of cortical direct current stimulation for the treatment of major depression*, International Journal of Neuropsychopharmacology, 11 (2008), pp. 249–254.
- [25] D. BRAESS, *Finite Elements: Theory, Fast Solvers, and Applications in Solid Mechanics*, Cambridge University Press, 3 ed., 2007.
- [26] V. BRODBECK, L. SPINELLI, A. M. LASCANO, M. WISSMEIER, M.-I. VARGAS, S. VULLIEMOZ, C. POLLO, K. SCHALLER, C. M. MICHEL, AND M. SEECK, *Electroencephalographic source imaging: a prospective study of 152 operated epileptic patients*, Brain, 134 (2011), pp. 2887–2897.
- [27] H. BUCHNER, M. FUCHS, H. A. WISCHMANN, O. DÖSSEL, I. LUDWIG, A. KNEPPER, AND P. BERG, *Source analysis of median nerve and finger stimulated somatosensory evoked potentials: Multichannel simultaneous recording of electric and magnetic fields combined with 3d-MR tomography*, Brain Topography, 6 (1994), pp. 299–310.
- [28] H. BUCHNER, G. KNOLL, M. FUCHS, A. RIENÄCKER, R. BECKMANN, M. WAGNER, J. SILNY, AND J. PESCH, *Inverse localization of electric dipole current sources in finite element models of the human head*, Electroencephalography and Clinical Neurophysiology, 102 (1997), pp. 267–278.
- [29] J. R. BUNCH, C. P. NIELSEN, AND D. C. SORESENSEN, *Rank-one modification of the symmetric eigenproblem*, Numerische Mathematik, 31 (1978), p. 31 – 48.
- [30] Y. BUSCHERMÖHLE, M. B. HÖLTERSINKEN, T. ERDBRÜGGER, J.-O. RADECKE, A. SPRENGER, T. R. SCHNEIDER, R. LENCER, J. GROSS, AND C. H. WOLTERS, *Comparing the performance of beamformer algorithms in estimating orientations of neural sources*, iScience, 27 (2024), p. 109150.
- [31] M. CARBONI, D. BRUNET, M. SEEGER, C. M. MICHEL, S. VULLIEMOZ, AND B. J. VORDERWÜLBECKE, *Linear distributed inverse solutions for interictal EEG source localization*, Clinical Neurophysiology, 133 (2022), pp. 58–67.
- [32] J. CATALÁN-AGUILAR, E. GONZÁLEZ-BONO, AND I. CANO-LÓPEZ, *Perceived stress in adults with epilepsy: A systematic review*, Neuroscience & Biobehavioral Reviews, 170 (2025), p. 106065.
- [33] W. C. CHEW, *A New Look at Reciprocity and Energy Conservation Theorems in Electromagnetics*, IEEE Transactions on Antennas and Propagation, 56 (2008), pp. 970–975.
- [34] C. COSMO, A. ZANDVAKILI, N. J. PETROSINO, Y. A. BERLOW, AND N. S. PHILIP, *Repetitive Transcranial Magnetic Stimulation for Treatment-Resistant Depression: Recent Critical Advances in Patient Care*, Current Treatment Options in Psychiatry, 8 (2021), pp. 47–63.
- [35] O. D. CREUTZFELDT, G. H. FROMM, AND H. KAPP, *Influence of transcortical d-c currents on cortical neuronal activity*, Experimental Neurology, 5 (1962), pp. 436–452.

- [36] E. CUARTAS, C. ACOSTA-MEDINA, G. CASTELLANOS-DOMINGUEZ, AND D. MANTINI, *A finite-difference solution for the EEG forward problem in inhomogeneous anisotropic media*, Brain Topography, 32 (2019).
- [37] I. CVX RESEARCH, *CVX: Matlab Software for Disciplined Convex Programming, version 2.0*. <https://cvxr.com/cvx>, Aug. 2012.
- [38] A. M. DALE, A. K. LIU, B. R. FISCHL, R. L. BUCKNER, J. W. BELLIVEAU, J. D. LEWINE, AND E. HALGREN, *Dynamic Statistical Parametric Mapping: Combining fMRI and MEG for High-Resolution Imaging of Cortical Activity*, Neuron, 26 (2000), pp. 55–67.
- [39] J. DE MUNCK AND M. PETERS, *A fast method to compute the potential in the multi-sphere model (EEG application)*, IEEE Transactions on Biomedical Engineering, 40 (1993), pp. 1166–1174.
- [40] J. DE MUNCK, B. VAN DIJK, AND H. SPEKREIJSE, *Mathematical dipoles are adequate to describe realistic generators of human brain activity*, IEEE Transactions on Biomedical Engineering, 35 (1988), pp. 960–966.
- [41] J. C. DE MUNCK, B. W. VAN DIJK, AND H. SPEKREIJSE, *An analytic method to determine the effect of source modeling errors on the apparent location and direction of biological sources*, Journal of Applied Physics, 63 (1988), pp. 944–956.
- [42] J. C. DE MUNCK, C. H. WOLTERS, AND M. CLERC, *EEG and MEG: forward modeling*, Cambridge University Press, 2012, p. 192–256.
- [43] J. DE TISI, G. S. BELL, J. L. PEACOCK, A. W. MCEVOY, W. F. HARKNESS, J. W. SANDER, AND J. S. DUNCAN, *The long-term outcome of adult epilepsy surgery, patterns of seizure remission, and relapse: a cohort study*, The Lancet, 378 (2011), pp. 1388–1395.
- [44] J. DIEUDONNE, *Foundations of Modern Analysis*, Pure and Applied Mathematics, Read Books, 2008.
- [45] J. P. DMOCHOWSKI, A. DATTA, M. BIKSON, Y. SU, AND L. C. PARRA, *Optimized multi-electrode stimulation increases focality and intensity at target*, Journal of Neural Engineering, 8 (2011), p. 046011.
- [46] J. P. DMOCHOWSKI, L. KOESSLER, A. M. NORCIA, M. BIKSON, AND L. C. PARRA, *Optimal use of EEG recordings to target active brain areas with transcranial electrical stimulation*, NeuroImage, 157 (2017), pp. 69–80.
- [47] M. DRAKAKI, C. MATHIESEN, H. R. SIEBNER, K. MADSEN, AND A. THIELSCHER, *Database of 25 validated coil models for electric field simulations for TMS*, Brain Stimulation, 15 (2022), pp. 697–706.
- [48] F. DRECHSLER, C. WOLTERS, T. DIERKES, H. SI, AND L. GRASEDYCK, *A full subtraction approach for finite element method based source analysis using constrained Delaunay tetrahedralisation*, NeuroImage, 46 (2009), pp. 1055–65.
- [49] R. E. ELLIOTT, A. MORSI, S. P. KALHORN, J. MARCUS, J. SELLIN, M. KANG, A. SILVERBERG, E. RIVERA, E. GELLER, C. CARLSON, O. DEVINSKY, AND W. K. DOYLE, *Vagus nerve stimulation in 436 consecutive patients with treatment-resistant epilepsy: Long-term outcomes and predictors of response*, Epilepsy & Behavior, 20 (2011), pp. 57–63.

- [50] J. ENGEL, S. WIEBE, J. FRENCH, M. SPERLING, P. WILLIAMSON, D. SPENCER, R. GUMNIT, C. ZAHN, E. WESTBROOK, AND B. ENOS, *Practice parameter: Temporal lobe and localized neocortical resections for epilepsy*, *Neurology*, 60 (2003), pp. 538–547.
- [51] D. A. ENGEMANN AND A. GRAMFORT, *Automated model selection in covariance estimation and spatial whitening of MEG and EEG signals*, *NeuroImage*, 108 (2015), pp. 328–342.
- [52] D. J. ENGLLOT, S. S. NAGARAJAN, B. S. IMBER, K. P. RAYGOR, S. M. HONMA, D. MIZUIRI, M. MANTLE, R. C. KNOWLTON, H. E. KIRSCH, AND E. F. CHANG, *Epileptogenic zone localization using magnetoencephalography predicts seizure freedom in epilepsy surgery*, *Epilepsia*, 56 (2015), pp. 949–958.
- [53] C. ENGWER, J. VORWERK, J. LUDEWIG, AND C. H. WOLTERS, *A discontinuous Galerkin method to solve the EEG forward problem using the subtraction approach*, *SIAM Journal on Scientific Computing*, 39 (2017), pp. B138–B164.
- [54] T. ERDBRÜGGER, M. ANTONAKAKIS, U. AYDIN, A. WESTHOFF, M. HÖLTERSHINKEN, J. VORWERK, S. SCHRADER, J.-O. RADECKE, S. WAGNER, T. MEDANI, M. C. PIASTRA, A. KHAN, F. NEUGEBAUER, S. PURSIAINEN, C. ENGWER, AND C. H. WOLTERS, *The University of Münster DUNEuro-based pipeline to create personalized head models with calibrated skull conductivity for EEG/MEG source analysis and optimized multi-channel tES*, Sept. 2024.
- [55] T.-R. ERDBRÜGGER, *CutFEM for EEG and MEG source analysis and optimized multi-channel transcranial direct current stimulation*, dissertation, University of Muenster, 2024.
- [56] T. ERDBRÜGGER, M. HÖLTERSHINKEN, J.-O. RADECKE, Y. BUSCHERMÖHLE, F. WALLOIS, S. PURSIAINEN, J. GROSS, R. LENCER, C. ENGWER, AND C. WOLTERS, *CutFEM-based MEG forward modeling improves source separability and sensitivity to quasi-radial sources: A somatosensory group study*, *Human Brain Mapping*, 45 (2024), p. e26810.
- [57] L. C. EVANS, *Partial Differential Equations: Second Edition*, 2010.
- [58] M. FERNÁNDEZ-CORAZZA, S. TUROVETS, P. LUU, E. ANDERSON, AND D. TUCKER, *Transcranial Electrical Neuromodulation Based on the Reciprocity Principle*, *Frontiers in Psychiatry*, Volume 7 - 2016 (2016).
- [59] M. FERNÁNDEZ-CORAZZA, S. TUROVETS, P. LUU, N. PRICE, C. H. MURAVCHIK, AND D. TUCKER, *Skull Modeling Effects in Conductivity Estimates Using Parametric Electrical Impedance Tomography*, *IEEE Transactions on Biomedical Engineering*, 65 (2018), pp. 1785–1797.
- [60] M. FERNÁNDEZ-CORAZZA, S. TUROVETS, AND C. H. MURAVCHIK, *Unification of optimal targeting methods in transcranial electrical stimulation*, *NeuroImage*, 209 (2020), p. 116403.
- [61] T. FERREE, K. ERIKSEN, AND D. TUCKER, *Regional head tissue conductivity estimation for improved EEG analysis*, *IEEE Transactions on Biomedical Engineering*, 47 (2000), pp. 1584–1592.
- [62] ———, *Correction to "Regional head tissue conductivity estimation for improved EEG analysis"*, *IEEE Transactions on Biomedical Engineering*, 48 (2001), pp. 742–.
- [63] K. M. FIEST, K. M. SAURO, S. WIEBE, S. B. PATTEN, C.-S. KWON, J. DYKEMAN, T. PRINGSHEIM, D. L. LORENZETTI, AND N. JETTÉ, *Prevalence and incidence of epilepsy*, *Neurology*, 88 (2017), pp. 296–303.

- [64] O. FORSTER, *Analysis 3: Maß- und Integrationstheorie, Integralsätze im \mathbb{R}^n und Anwendungen*, Aufbaukurs Mathematik, Vieweg+Teubner Verlag, 2012.
- [65] J. T. FRANCIS, B. J. GLUCKMAN, AND S. J. SCHIFF, *Sensitivity of Neurons to Weak Electric Fields*, Journal of Neuroscience, 23 (2003), pp. 7255–7261.
- [66] A. FRANK, *Sensitivity of Optimization in Transcranial Direct Current Stimulation to Electrode Modeling*, Master’s thesis, University of Münster, 2022.
- [67] M. FUCHS, M. WAGNER, T. KÖHLER, AND H.-A. WISCHMANN, *Linear and Nonlinear Current Density Reconstructions*, Journal of Clinical Neurophysiology, 16 (1999).
- [68] M. FUCHS, M. WAGNER, H.-A. WISCHMANN, T. KÖHLER, A. THEISSEN, R. DRENCKHAHN, AND H. BUCHNER, *Improving source reconstructions by combining bioelectric and biomagnetic data*, Electroencephalography and Clinical Neurophysiology, 107 (1998), pp. 93–111.
- [69] F. GALAZ PRIETO, M. SAMAVAKI, AND S. PURSIAINEN, *Lattice layout and optimizer effect analysis for generating optimal transcranial electrical stimulation (tES) montages through the metaheuristic L1L1 method*, Frontiers in Human Neuroscience, Volume 18 - 2024 (2024).
- [70] N. GAO, S. A. ZHU, AND B. HE, *A new magnetic resonance electrical impedance tomography (MREIT) algorithm: the RSM-MREIT algorithm with applications to estimation of human head conductivity*, Physics in Medicine & Biology, 51 (2006), p. 3067.
- [71] D. GESELOWITZ, *On the magnetic field generated outside an inhomogeneous volume conductor by internal current sources*, IEEE Transactions on Magnetism, 6 (1970), pp. 346–347.
- [72] C. GEUZAIN AND J.-F. REMACLE, *Gmsh: A 3-d finite element mesh generator with built-in pre- and post-processing facilities*, International Journal for Numerical Methods in Engineering, 79 (2009), pp. 1309–1331.
- [73] M. GNEWUCH AND S. SAUTER, *Boundary integral equations for second order elliptic boundary value problems*, (1999).
- [74] D. M. GOLDENHOLZ, S. P. AHLFORS, M. S. HÄMÄLÄINEN, D. SHARON, M. ISHITOBI, L. M. VAINA, AND S. M. STUFFLEBEAM, *Mapping the signal-to-noise-ratios of cortical sources in magnetoencephalography and electroencephalography*, Human Brain Mapping, 30 (2009), pp. 1077–1086.
- [75] R. GRAGLIA, *On the numerical integration of the linear shape functions times the 3-d Green’s function or its gradient on a plane triangle*, IEEE Transactions on Antennas and Propagation, 41 (1993), pp. 1448–1455.
- [76] M. GRANT AND S. BOYD, *Graph implementations for nonsmooth convex programs*, in Recent Advances in Learning and Control, V. Blondel, S. Boyd, and H. Kimura, eds., Lecture Notes in Control and Information Sciences, Springer-Verlag Limited, 2008, pp. 95–110. http://stanford.edu/~boyd/graph_dcp.html.
- [77] R. GRECH, T. CASSAR, J. MUSCAT, K. CAMILLERI, S. FABRI, M. ZERVAKIS, P. XANTHOPOULOS, V. SAKKALIS, AND B. VANRUMSTE, *Review on solving the inverse problem in EEG source analysis*, Journal of neuroengineering and rehabilitation, 5 (2008), p. 25.

- [78] R. GREENBLATT, A. OSSADTCHI, AND M. PFLIEGER, *Local Linear Estimators for the Bioelectromagnetic Inverse Problem*, Signal Processing, IEEE Transactions on, 53 (2005), pp. 3403 – 3412.
- [79] T. GRONOTTE, *Impact of Grey Matter Conductivity Anisotropy on EEG and MEG Source Estimation (forthcoming)*, Master’s thesis, University of Muenster, 2025.
- [80] J. GROSS, *Magnetoencephalography in Cognitive Neuroscience: A Primer*, Neuron, 104 (2019), pp. 189–204.
- [81] J. GROSS, M. JUNGHÖFER, AND C. WOLTERS, *Bioelectromagnetism in Human Brain Research: New Applications, New Questions*, The Neuroscientist, 29 (2023), pp. 62–77. PMID: 34873945.
- [82] D. GÜLLMAR, J. HAUEISEN, AND J. R. REICHENBACH, *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 (2010), pp. 145–163.
- [83] M. HÄMÄLÄINEN, R. HARI, R. J. ILMONIEMI, J. KNUUTILA, AND O. V. LOUNASMAA, *Magnetoencephalography—theory, instrumentation, and applications to noninvasive studies of the working human brain*, Rev. Mod. Phys., 65 (1993), pp. 413–497.
- [84] M. HANKE, B. HARRACH, AND N. HYVÖNEN, *Justification of point electrode models in electrical impedance tomography*, Mathematical Models and Methods in Applied Sciences, 21 (2011), pp. 1395–1413.
- [85] M. HANKE, N. HYVÖNEN, AND S. REUSSWIG, *Convex backscattering support in electric impedance tomography*, Numerische Mathematik, 117 (2011), pp. 373–396.
- [86] A. HANRATH, *Finite Element Representation of the EEG Forward Problem with Multipole Expansion*, PhD thesis, RWTH Aachen, 2018.
- [87] J. HAUEISEN, D. TUCH, C. RAMON, P. SCHIMPF, V. WEDEEN, J. GEORGE, AND J. BELLIVEAU, *The Influence of Brain Tissue Anisotropy on Human EEG and MEG*, NeuroImage, 15 (2002), pp. 159–166.
- [88] R. M. HEIDEMANN, D. A. PORTER, A. ANWANDER, T. FEIWEIER, K. HEBERLEIN, T. R. KNÖSCHE, AND R. TURNER, *Diffusion imaging in humans at 7T using readout-segmented EPI and GRAPPA*, Magnetic Resonance in Medicine, 64 (2010), pp. 9–14.
- [89] H. HELMHOLTZ, *Ueber einige Gesetze der Vertheilung elektrischer Ströme in körperlichen Leitern mit Anwendung auf die thierisch-elektrischen Versuche*, Annalen der Physik, 165 (1853), pp. 211–233.
- [90] M. HOELTERSHINKEN, *Maltehoel/local_subtraction_paper_code: Release 3.0*, Feb. 2025.
- [91] M. B. HÖLTERSHINKEN, P. LANGE, T. ERDBRÜGGER, Y. BUSCHERMÖHLE, F. WALLOIS, A. BUYX, S. PURSIAINEN, J. VORWERK, C. ENGWER, AND C. H. WOLTERS, *The Local Subtraction Approach for EEG and MEG Forward Modeling*, SIAM Journal on Scientific Computing, 47 (2025), pp. B160–B189.
- [92] M.-X. HUANG, T. SONG, D. J. HAGLER, I. PODGORNÝ, V. JOUSMAKI, L. CUI, K. GAA, D. L. HARRINGTON, A. M. DALE, R. R. LEE, J. ELMAN, AND E. HALGREN, *A novel integrated MEG and EEG analysis method for dipolar sources*, NeuroImage, 37 (2007), pp. 731–748.

- [93] F. HUMMEL, P. CELNIK, P. GIRAUX, A. FLOEL, W.-H. WU, C. GERLOFF, AND L. G. COHEN, *Effects of non-invasive cortical stimulation on skilled motor function in chronic stroke*, Brain, 128 (2005), pp. 490–499.
- [94] M. HÄMÄLÄINEN AND R. ILMONIEMI, *Interpreting magnetic fields of the brain: minimum-norm estimates*, Medical & biological engineering & computing, 32 (1994), pp. 35–42.
- [95] M. B. HÖLTERSHINKEN, T. ERDBRÜGGER, AND C. H. WOLTERS, *sLORETA is equivalent to single dipole scanning*, 2024.
- [96] M. B. HÖLTERSHINKEN, P. LANGE, T. ERDBRÜGGER, Y. BUSCHERMÖHLE, F. WALLOIS, A. BUYX, S. PURSIAINEN, J. VORWERK, C. ENGWER, AND C. H. WOLTERS, *Data used for validation of the Local Subtraction Approach*, June 2024.
- [97] ———, *Maltehoel/duneuro_local_subtraction: Implementation as described in SIAM SISC paper*, June 2024.
- [98] R. ILMONIEMI, J. RUOHONEN, AND J. KARHU, *Transcranial magnetic stimulation - A new tool for functional imaging of the brain*, Critical reviews in biomedical engineering, 27 (1999), pp. 241–84.
- [99] M. JENKINSON, P. BANNISTER, M. BRADY, AND S. SMITH, *Improved Optimization for the Robust and Accurate Linear Registration and Motion Correction of Brain Images*, NeuroImage, 17 (2002), pp. 825–841.
- [100] M. JENKINSON, C. F. BECKMANN, T. E. BEHRENS, M. W. WOOLRICH, AND S. M. SMITH, *FSL*, NeuroImage, 62 (2012), pp. 782–790. 20 YEARS OF fMRI.
- [101] K. JERBI, S. BAILLET, J. MOSHER, G. NOLTE, L. GARNERO, AND R. LEAHY, *Localization of realistic cortical activity in MEG using current multipoles*, NeuroImage, 22 (2004), pp. 779–793.
- [102] N. KANE, J. ACHARYA, S. BENICZKY, L. CABOCLO, S. FINNIGAN, P. W. KAPLAN, H. SHIBASAKI, R. PRESSLER, AND M. J. VAN PUTTEN, *A revised glossary of terms most commonly used by clinical electroencephalographers and updated proposal for the report format of the EEG findings. Revision 2017*, Clinical Neurophysiology Practice, 2 (2017), pp. 170–185.
- [103] F. KARPUSZ, *The Complete Electrode Model in Bioelectromagnetism: Analysis, Implementation and Evaluation*, Master’s thesis, University of Muenster, 2025.
- [104] E. KAUFMANN, M. HORDT, M. LAUSEKER, U. PALM, AND S. NOACHTAR, *Acute effects of spaced cathodal transcranial direct current stimulation in drug resistant focal epilepsies*, Clinical Neurophysiology, 132 (2021), pp. 1444–1451.
- [105] A. KHAN, M. ANTONAKAKIS, S. SUNTRUP-KRUEGER, R. LENCER, M. A. NITSCHKE, W. PAULUS, J. GROSS, AND C. H. WOLTERS, *Can individually targeted and optimized multi-channel tDCS outperform standard bipolar tDCS in stimulating the primary somatosensory cortex?*, Brain Stimulation: Basic, Translational, and Clinical Research in Neuromodulation, 16 (2023), pp. 1–16.
- [106] A. KHAN, M. ANTONAKAKIS, N. VOGENAUER, J. HAUEISEN, AND C. H. WOLTERS, *Individually optimized multi-channel tDCS for targeting somatosensory cortex*, Clinical Neurophysiology, 134 (2022), pp. 9–26.

- [107] T. D. KRIEG, F. S. SALINAS, S. NARAYANA, P. T. FOX, AND D. J. MOGUL, *Computational and experimental analysis of TMS-induced electric field vectors critical to neuronal activation*, Journal of Neural Engineering, 12 (2015), p. 046014.
- [108] P. KÜPPER, *Combined EEG and MEG for improving source analysis in patients with focal epilepsy*, Master's thesis, Beuth Hochschule für Technik Berlin, 2012.
- [109] P. KWAN, A. ARZIMANOGLU, A. T. BERG, M. J. BRODIE, W. ALLEN HAUSER, G. MATHERN, S. L. MOSHÉ, E. PERUCCA, S. WIEBE, AND J. FRENCH, *Definition of drug resistant epilepsy: Consensus proposal by the ad hoc Task Force of the ILAE Commission on Therapeutic Strategies*, Epilepsia, 51 (2010), pp. 1069–1077.
- [110] J. KYBIC, M. CLERC, T. ABBOD, O. FAUGERAS, R. KERIVEN, AND T. PAPADOPOULOU, *A common formalism for the integral formulations of the forward EEG problem*, IEEE Transactions on Medical Imaging, 24 (2005), pp. 12–28.
- [111] J. KYBIC, M. CLERC, O. FAUGERAS, R. KERIVEN, AND T. PAPADOPOULOU, *Fast multipole acceleration of the MEG/EEG boundary element method*, Physics in Medicine & Biology, 50 (2005), p. 4695.
- [112] I. LAAKSO, S. TANAKA, S. KOYAMA, V. DE SANTIS, AND A. HIRATA, *Inter-subject Variability in Electric Fields of Motor Cortical tDCS*, Brain Stimulation, 8 (2015), pp. 906–913.
- [113] B. LABEIT, E. MICHOU, M. TRAPL-GRUNDSCHÖBER, S. SUNTRUP-KRUEGER, P. MUHLE, P. M. BATH, AND R. DZIEWAS, *Dysphagia after stroke: research advances in treatment interventions*, The Lancet Neurology, 23 (2024), pp. 418–428.
- [114] J. LAHTINEN, P. RONNI, N. P. SUBRAMANIAM, A. KOULOURI, C. WOLTERS, AND S. PURSIAINEN, *Standardized Kalman filtering for dynamical source localization of concurrent subcortical and cortical brain activity*, Clinical Neurophysiology, 168 (2024), pp. 15–24.
- [115] G. LANTZ, L. SPINELLI, M. SEECK, R. G. DE PERALTA MENENDEZ, C. C. SOTTAS, AND C. M. MICHEL, *Propagation of Interictal Epileptiform Activity Can Lead to Erroneous Source Localizations: A 128-Channel EEG Mapping Study*, Journal of Clinical Neurophysiology, 20 (2003).
- [116] S. LEW, C. WOLTERS, T. DIERKES, C. RÖER, AND R. MACLEOD, *Accuracy and run-time comparison for different potential approaches and iterative solvers in finite element method based EEG source analysis*, Applied Numerical Mathematics, 59 (2009), pp. 1970–1988.
- [117] J. LI, K. WANG, S. ZHU, AND B. HE, *Effects of holes on EEG forward solutions using a realistic geometry head model*, Journal of Neural Engineering, 4 (2007), p. 197.
- [118] L. M. LI, K. UEHARA, AND T. HANAKAWA, *The contribution of interindividual factors to variability of response in transcranial direct current stimulation studies*, Frontiers in Cellular Neuroscience, Volume 9 - 2015 (2015).
- [119] F.-H. LIN, T. WITZEL, S. P. AHLFORS, S. M. STUFFLEBEAM, J. W. BELLIVEAU, AND M. S. HÄMÄLÄINEN, *Assessing and improving the spatial accuracy in MEG source localization by depth-weighted minimum-norm estimates*, NeuroImage, 31 (2006), pp. 160–171.
- [120] J. LIONS AND E. MAGENES, *Non-Homogeneous Boundary Value Problems and Applications*, 1972.

- [121] A. LIU, M. VÖRÖSLAKOS, G. KRONBERG, S. HENIN, M. R. KRAUSE, Y. HUANG, A. OPITZ, A. MEHTA, C. C. PACK, B. KREKELBERG, A. BERÉNYI, L. C. PARRA, L. MELLONI, O. DEVINSKY, AND G. BUZSÁKI, *Immediate neurophysiological effects of transcranial electrical stimulation*, Nature Communications, 9 (2018), p. 5092.
- [122] V. LÓPEZ-ALONSO, B. CHEERAN, D. RÍO-RODRÍGUEZ, AND M. FERNÁNDEZ-DEL OLMO, *Inter-individual Variability in Response to Non-invasive Brain Stimulation Paradigms*, Brain Stimulation: Basic, Translational, and Clinical Research in Neuromodulation, 7 (2014), pp. 372–380.
- [123] K. H. MADSEN, M. DRAKAKI, AND A. THIELSCHER, *Electric Field Models of Transcranial Magnetic Stimulation Coils with Arbitrary Geometries: Reconstruction from Incomplete Magnetic Field Measurements*, 2021.
- [124] S. N. MAKAROV, M. HÄMÄLÄINEN, Y. OKADA, G. M. NOETSCHER, J. AHVENINEN, AND A. NUMMENMAA, *Boundary element fast multipole method for enhanced modeling of neurophysiological recordings*, IEEE Transactions on Biomedical Engineering, 68 (2021), pp. 308–318.
- [125] J. MALMIVUO, *Principle of reciprocity solves the most important problems in bioimpedance and in general in bioelectromagnetism*, Journal of Physics Conference Series, 224 (2011), pp. 4–7.
- [126] G. MARIN, C. GUERIN, S. BAILLET, L. GARNERO, AND G. MEUNIER, *Influence of skull anisotropy for the forward and inverse problem in EEG: Simulation studies using FEM on realistic head models*, Human Brain Mapping, 6 (1998), pp. 250–269.
- [127] S. MASOUD MOOSAVI BASRI, H. AL-NASHASH, AND H. MIR, *Inverse Problem for M/EEG Source Localization: A Review*, IEEE Sensors Journal, 25 (2025), pp. 2103–2124.
- [128] H. MCCANN, G. PISANO, AND L. BELTRACHINI, *Variation in Reported Human Head Tissue Electrical Conductivity Values*, Brain Topography, 32 (2019), pp. 825–858.
- [129] T. MEDANI, J. GARCIA-PRIETO, F. TADEL, M. ANTONAKAKIS, T. ERDBRÜGGER, M. HÖLTERSINKEN, W. MEAD, S. SCHRADER, A. JOSHI, C. ENGWER, C. H. WOLTERS, J. C. MOSHER, AND R. M. LEAHY, *Brainstorm-DUNEuro: An integrated and user-friendly Finite Element Method for modeling electromagnetic brain activity*, NeuroImage, 267 (2023), p. 119851.
- [130] T. MEDANI, D. LAUTRU, D. SCHWARTZ, Z. REN, AND G. SOU, *FEM Method for the EEG Forward Problem and Improvement Based on Modification of the Saint Venant’s Method*, Progress in Electromagnetics Research, 153 (2015), pp. 11–22.
- [131] D. R. MERRILL, M. BIKSON, AND J. G. JEFFERYS, *Electrical stimulation of excitable tissue: design of efficacious and safe protocols*, Journal of Neuroscience Methods, 141 (2005), pp. 171–198.
- [132] M. MERTENS AND B. LÜTKENHÖNER, *Efficient neuromagnetic determination of landmarks in the somatosensory cortex*, Clinical Neurophysiology, 111 (2000), pp. 1478–1487.
- [133] T. MIINALAINEN, A. REZAEI, D. US, A. NÜSSING, C. ENGWER, C. WOLTERS, AND S. PURSIAINEN, *A realistic, accurate and fast source modeling approach for the EEG forward problem*, NeuroImage, 184 (2018).

- [134] M. F. MOHD ZULKIFLY, A. LEHR, D. VAN DE VELDEN, A. KHAN, N. K. FOCKE, C. H. WOLTERS, AND W. PAULUS, *Directionality of the injected current targeting the P20/N20 source determines the efficacy of 140 Hz transcranial alternating current stimulation (tACS)-induced aftereffects in the somatosensory cortex*, PLOS ONE, 17 (2022), pp. 1–24.
- [135] A. MOISEEV, J. M. GASPAR, J. A. SCHNEIDER, AND A. T. HERDMAN, *Application of multi-source minimum variance beamformers for reconstruction of correlated neural activity*, NeuroImage, 58 (2011), pp. 481–496.
- [136] N. MÄKELÄ, M. STENROOS, J. SARVAS, AND R. J. ILMONIEMI, *Truncated RAP-MUSIC (TRAP-MUSIC) for MEG and EEG source localization*, NeuroImage, 167 (2018), pp. 73–83.
- [137] F. NEUGEBAUER, G. MÖDDEL, S. RAMPP, M. BURGER, AND C. H. WOLTERS, *The Effect of Head Model Simplification on Beamformer Source Localization*, Frontiers in Neuroscience, Volume 11 - 2017 (2017).
- [138] J. D. NIELSEN, K. H. MADSEN, O. PUONTI, H. R. SIEBNER, C. BAUER, C. G. MADSEN, G. B. SATURNINO, AND A. THIELSCHER, *Automatic skull segmentation from MR images for realistic volume conductor models of the head: Assessment of the state-of-the-art*, NeuroImage, 174 (2018), pp. 587–598.
- [139] M. A. NITSCHKE AND W. PAULUS, *Excitability changes induced in the human motor cortex by weak transcranial direct current stimulation*, The Journal of Physiology, 527 (2000), pp. 633–639.
- [140] G. NOLTE, *The magnetic lead field theorem in the quasi-static approximation and its use for magnetoencephalography forward calculation in realistic volume conductors*, Physics in Medicine & Biology, 48 (2003), p. 3637.
- [141] W. NOLTING, *Grundkurs Theoretische Physik 3: Elektrodynamik*, 01 2007.
- [142] M. NOWELL, R. RODIONOV, B. DIEHL, T. WEHNER, G. ZOMBORI, J. KINGHORN, S. OURSELIN, J. DUNCAN, A. MISEROCCHI, AND A. MCEVOY, *A Novel Method for Implementation of Frameless StereoEEG in Epilepsy Surgery*, Operative Neurosurgery, 10 (2014).
- [143] A. NÜSSING, *Fitted and unfitted finite element methods for solving the EEG forward problem*, dissertation, University of Muenster, 2018.
- [144] G. NUÑEZ PONASSO, W. A. WARTMAN, R. C. MCSWEENEY, P. LAI, J. HAUEISEN, B. MAESS, T. R. KNÖSCHE, K. WEISE, G. M. NOETSCHER, T. RAIJ, AND S. N. MAKAROFF, *Improving EEG Forward Modeling Using High-Resolution Five-Layer BEM-FMM Head Models: Effect on Source Reconstruction Accuracy*, Bioengineering, 11 (2024).
- [145] E. OLIVI, T. PAPADOPOULOU, AND M. CLERC, *Handling white-matter anisotropy in BEM for the EEG forward problem*, in 2011 IEEE International Symposium on Biomedical Imaging: From Nano to Macro, 2011, pp. 799–802.
- [146] R. OOSTENVELD, P. FRIES, E. MARIS, AND J.-M. SCHOFFELEN, *FieldTrip: Open Source Software for Advanced Analysis of MEG, EEG, and Invasive Electrophysiological Data*, Computational Intelligence and Neuroscience, 2011 (2011), p. 156869.

- [147] R. OOSTENVELD AND T. F. OOSTENDORP, *Validating the boundary element method for forward and inverse eeg computations in the presence of a hole in the skull*, Human Brain Mapping, 17 (2002), pp. 179–192.
- [148] L. J. O'DONNELL AND C.-F. WESTIN, *An introduction to diffusion tensor image analysis*, Neurosurgery Clinics of North America, 22 (2011), pp. 185–196. Functional Imaging.
- [149] R. D. PASCUAL-MARQUI, *Standardized low-resolution brain electromagnetic tomography (sLORETA): technical details.*, Methods and findings in experimental and clinical pharmacology, 24 Suppl D (2002), pp. 5–12.
- [150] ———, *Discrete, 3D distributed, linear imaging methods of electric neuronal activity. Part 1: exact, zero error localization*, 2007.
- [151] M. C. PIASTRA, *New Finite Element Methods for Solving the MEG and the Combined MEG/EEG Forward Problem*, dissertation, University of Muenster, 2019.
- [152] M. C. PIASTRA, A. NÜSSING, J. VORWERK, H. BORNFLETH, R. OOSTENVELD, C. ENGWER, AND C. H. WOLTERS, *The discontinuous Galerkin finite element method for solving the MEG and the combined MEG/EEG forward problem*, Frontiers in Neuroscience, 12 (2018).
- [153] M. C. PIASTRA, A. NÜSSING, J. VORWERK, M. CLERC, C. ENGWER, AND C. H. WOLTERS, *A comprehensive study on electroencephalography and magnetoencephalography sensitivity to cortical and subcortical sources*, Human Brain Mapping, 42 (2021), pp. 978–992.
- [154] M. C. PIASTRA, S. SCHRADER, A. NÜSSING, M. ANTONAKAKIS, T. MEDANI, A. WOLLBRINK, C. ENGWER, AND C. H. WOLTERS, *The WWU DUNEuro reference data set for combined EEG/MEG source analysis*, June 2020.
- [155] M.-C. PICOT, A. JAUSSENT, D. NEVEU, P. KAHANE, A. CRESPEL, P. GELISSE, E. HIRSCH, P. DERAMBURE, S. DUPONT, E. LANDRÉ, F. CHASSOUX, L. VALTON, J.-P. VIGNAL, C. MARCHAL, C. LAMY, F. SEMAH, A. BIRABEN, A. ARZIMANOGLU, J. PETIT, P. THOMAS, V. MACIOCE, P. DUJOLS, AND P. RYVLIN, *Cost-effectiveness analysis of epilepsy surgery in a controlled cohort of adult patients with intractable partial epilepsy: A 5-year follow-up study*, Epilepsia, 57 (2016), pp. 1669–1679.
- [156] A. PILLAIN, L. RAHMOUNI, AND F. ANDRIULLI, *Handling anisotropic conductivities in the EEG forward problem with a symmetric formulation*, Physics in Medicine & Biology, 64 (2019), p. 035022.
- [157] C. PLUMMER, S. J. VOGGIN, W. P. WOODS, M. A. MURPHY, M. J. COOK, AND D. T. J. LILEY, *Interictal and ictal source localization for epilepsy surgery using high-density EEG with MEG: a prospective long-term study*, Brain, 142 (2019), pp. 932–951.
- [158] O. PUONTI, K. VAN LEEMPUT, G. B. SATURNINO, H. R. SIEBNER, K. H. MADSEN, AND A. THIELSCHER, *Accurate and robust whole-head segmentation from magnetic resonance images for individualized head modeling*, NeuroImage, 219 (2020), p. 117044.
- [159] S. PURSIAINEN, *Raviart–Thomas-type sources adapted to applied EEG and MEG: implementation and results*, Inverse Problems, 28 (2012), p. 065013.

- [160] S. PURSIAINEN, F. LUCKA, AND C. H. WOLTERS, *Complete electrode model in EEG: relationship and differences to the point electrode model*, Physics in Medicine & Biology, 57 (2012), p. 999.
- [161] S. PURSIAINEN, A. SORRENTINO, C. CAMPI, AND M. PIANA, *Forward simulation and inverse dipole localization with the lowest order Raviart—Thomas elements for electroencephalography*, Inverse Problems, 27 (2011), p. 045003.
- [162] S. PURSIAINEN, J. VORWERK, AND C. WOLTERS, *Electroencephalography (EEG) Forward Modeling via $H(\text{div})$ Finite Element Sources with Focal Interpolation*, Physics in medicine and biology, 61 (2016).
- [163] T. RADMAN, R. L. RAMOS, J. C. BRUMBERG, AND M. BIKSON, *Role of cortical cell type and morphology in subthreshold and suprathreshold uniform electric field stimulation in vitro*, Brain Stimulation, 2 (2009), pp. 215–228.e3.
- [164] S. RAMPP, H. STEFAN, X. WU, M. KALTENHÄUSER, B. MAESS, F. C. SCHMITT, C. H. WOLTERS, H. HAMER, B. S. KASPER, S. SCHWAB, A. DOERFLER, I. BLÜMCKE, K. RÖSSLER, AND M. BUCHFELDER, *Magnetoencephalography for epileptic focus localization in a series of 1000 cases*, Brain, 142 (2019), pp. 3059–3071.
- [165] T. REED AND R. COHEN KADOSH, *Transcranial electrical stimulation (tES) mechanisms and its effects on cortical excitability and connectivity*, Journal of Inherited Metabolic Disease, 41 (2018), pp. 1123–1130.
- [166] S. REZAKHANI, M. AMIRI, S. WECKHUYSEN, AND G. A. KELIRIS, *Therapeutic efficacy of seizure onset zone-targeting high-definition cathodal tDCS in patients with drug-resistant focal epilepsy*, Clinical Neurophysiology, 136 (2022), pp. 219–227.
- [167] F. ROSENOW AND H. LÜDERS, *Presurgical evaluation of epilepsy*, Brain, 124 (2001), pp. 1683–1700.
- [168] S. M. ROSS, *Introduction to Probability and Statistics for Engineers and Scientists (Fifth Edition)*, Academic Press, Boston, fifth edition ed., 2014.
- [169] B. J. ROTH, L. G. COHEN, M. HALLETT, W. FRIAUF, AND P. J. BASSER, *A theoretical calculation of the electric field induced by magnetic stimulation of a peripheral nerve*, Muscle & Nerve, 13 (1990), pp. 734–741.
- [170] W. RUDIN, *Functional Analysis*, McGraw-Hill, 2 ed., 1991.
- [171] M. RULLMANN, A. ANWANDER, M. DANNHAUER, S. WARFIELD, F. DUFFY, AND C. WOLTERS, *EEG source analysis of epileptiform activity using a 1 mm anisotropic hexahedra finite element head model*, NeuroImage, 44 (2009), pp. 399–410.
- [172] L. RUTHOTTO, H. KUGEL, J. OLESCH, B. FISCHER, J. MODERSITZKI, M. BURGER, AND C. H. WOLTERS, *Diffeomorphic susceptibility artifact correction of diffusion-weighted magnetic resonance images*, Physics in Medicine & Biology, 57 (2012), p. 5715.
- [173] Y. SAAD, *Iterative Methods for Sparse Linear Systems*, 01 2003.
- [174] ———, *Numerical Methods for Large Eigenvalue Problems*, Society for Industrial and Applied Mathematics, 2011.
- [175] V. SALANOVA, T. WITT, R. WORTH, T. R. HENRY, R. E. GROSS, J. M. NAZZARO, D. LABAR, M. R. SPERLING, A. SHARAN, E. SANDOK, A. HANDFORTH, J. M. STERN,

- S. CHUNG, J. M. HENDERSON, J. FRENCH, G. BALTUCH, W. E. ROSENFELD, P. GARCIA, N. M. BARBARO, N. B. FOUNTAIN, W. J. ELIAS, R. R. GOODMAN, J. R. POLLARD, A. I. TRÖSTER, C. P. IRWIN, K. LAMBRECHT, N. GRAVES, R. FISHER, F. THE SANTE STUDY GROUP, C. EPSTEIN, J. KING, I. OSORIO, R. BAKAY, AND D. BERGEN, *Long-term efficacy and safety of thalamic stimulation for drug-resistant partial epilepsy*, *Neurology*, 84 (2015), pp. 1017–1025.
- [176] O. SANDER, *DUNE — The Distributed and Unified Numerics Environment*, Springer International Publishing, 2020.
- [177] D. T. SANDWELL, *Biharmonic spline interpolation of GEOS-3 and SEASAT altimeter data*, *Geophysical Research Letters*, 14 (1987), pp. 139–142.
- [178] J. SARVAS, *Basic mathematical and electromagnetic concepts of the biomagnetic inverse problem*, *Physics in Medicine and Biology*, 32 (1987), pp. 11–22.
- [179] G. B. SATURNINO, A. THIELSCHER, K. H. MADSEN, T. R. KNÖSCHE, AND K. WEISE, *A principled approach to conductivity uncertainty analysis in electric field calculations*, *NeuroImage*, 188 (2019), pp. 821–834.
- [180] I. E. SCHEFFER, S. BERKOVIC, G. CAPOVILLA, M. B. CONNOLLY, J. FRENCH, L. GUILHOTO, E. HIRSCH, S. JAIN, G. W. MATHERN, S. L. MOSHÉ, D. R. NORDLI, E. PERUCCA, T. TOMSON, S. WIEBE, Y.-H. ZHANG, AND S. M. ZUBERI, *Ilae classification of the epilepsies: Position paper of the ilae commission for classification and terminology*, *Epilepsia*, 58 (2017), pp. 512–521.
- [181] M. SCHERG, P. BERG, N. NAKASATO, AND S. BENICZKY, *Taking the EEG Back Into the Brain: The Power of Multiple Discrete Sources*, *Frontiers in Neurology*, Volume 10 - 2019 (2019).
- [182] P. SCHIMPF, C. RAMON, AND J. HAUEISEN, *Dipole models for the EEG and MEG*, *IEEE Transactions on Biomedical Engineering*, 49 (2002), pp. 409–418.
- [183] C. SCHMIDT, S. WAGNER, M. BURGER, U. V. RIENEN, AND C. H. WOLTERS, *Impact of uncertain head tissue conductivity in the optimization of transcranial direct current stimulation for an auditory target*, *Journal of Neural Engineering*, 12 (2015), p. 046028.
- [184] J.-M. SCHOFFELEN AND J. GROSS, *Source connectivity analysis with MEG and EEG*, *Human Brain Mapping*, 30 (2009), pp. 1857–1865.
- [185] S. SCHRADER, M. ANTONAKAKIS, S. RAMPP, C. ENGWER, AND C. H. WOLTERS, *A novel method for calibrating head models to account for variability in conductivity and its evaluation in a sphere model*, *Physics in Medicine & Biology*, 65 (2020), p. 245043.
- [186] S. SCHRADER, A. WESTHOFF, M. C. PIASTRA, T. MIINALAINEN, S. PURSIAINEN, J. VORWERK, H. BRINCK, C. H. WOLTERS, AND C. ENGWER, *Duneuro—a software toolbox for forward modeling in bioelectromagnetism*, *PLOS ONE*, 16 (2021), pp. 1–21.
- [187] A. SCHULZE-BONHAGE, M. HIRSCH, S. KNAKE, E. KAUFMANN, J. KEGELE, M. RADEMACHER, K. VONCK, V. A. COENEN, M. GLASER, C. JENKNER, Y. WINTER, S. GROPPA, AND E. S. GROUP, *Focal Cortex Stimulation With a Novel Implantable Device and Antiseizure Outcomes in 2 Prospective Multicenter Single-Arm Trials*, *JAMA Neurology*, 80 (2023), pp. 588–596.

- [188] A. SCHULZE-BONHAGE, M. HIRSCH, S. KNAKE, A. MERTENS, M. RADEMACHER, E. KAUFMANN, J. KEGELE, C. JENKNER, V. COENEN, M. GLASER, S. GROPPA, Y. WINTER, AND FOR THE EASEE STUDY GROUP, *Two-year outcomes of epicranial focal cortex stimulation in pharmacoresistant focal epilepsy*, *Epilepsia*, n/a.
- [189] F. SEGONNE, J. PACHECO, AND B. FISCHL, *Geometrically Accurate Topology-Correction of Cortical Surfaces Using Nonseparating Loops*, *IEEE Transactions on Medical Imaging*, 26 (2007), pp. 518–529.
- [190] K. SEKIHARA AND S. NAGARAJAN, *Adaptive Spatial Filters for Electromagnetic Brain Imaging*, 01 2008.
- [191] K. SEKIHARA, S. NAGARAJAN, D. POEPEL, AND A. MARANTZ, *Asymptotic SNR of scalar and vector minimum-variance beamformers for neuromagnetic source reconstruction*, *IEEE Transactions on Biomedical Engineering*, 51 (2004), pp. 1726–1734.
- [192] K. SEKIHARA, M. SAHANI, AND S. S. NAGARAJAN, *Localization bias and spatial resolution of adaptive and non-adaptive spatial filters for MEG source reconstruction*, *NeuroImage*, 25 (2005), pp. 1056–1067.
- [193] F. SEMAH, M.-C. PICOT, C. ADAM, D. BROGLIN, A. ARZIMANOGLU, B. BAZIN, D. CAVALCANTI, AND M. BAULAC, *Is the underlying cause of epilepsy a major prognostic factor for recurrence?*, *Neurology*, 51 (1998), pp. 1256–1262.
- [194] J. S. SHIMONY, R. C. MCKINSTRY, E. AKBUDAK, J. A. ARONOVITZ, A. Z. SNYDER, N. F. LORI, T. S. CULL, AND T. E. CONTURO, *Quantitative Diffusion-Tensor Anisotropy Brain MR Imaging: Normative Human Data and Anatomic Analysis*, *Radiology*, 212 (1999), pp. 770–784. PMID: 10478246.
- [195] N. SINHA, J. DAUWELS, M. KAISER, S. S. CASH, M. BRANDON WESTOVER, Y. WANG, AND P. N. TAYLOR, *Predicting neurosurgical outcomes in focal epilepsy patients using computational modelling*, *Brain*, 140 (2016), pp. 319–332.
- [196] E. SOMERSALO, M. CHENEY, AND D. ISAACSON, *Existence and Uniqueness for Electrode Models for Electric Current Computed Tomography*, *SIAM Journal on Applied Mathematics*, 52 (1992), pp. 1023–1040.
- [197] M. P. SP VAN DEN BROEK, H ZHOU, *Computation of neuromagnetic fields using finite-element method and Biot-Savart law*, *Meg. Biol. Eng. Comput.*, 34 (1996), pp. 21–26.
- [198] A. STROUD, *Approximate Calculation of Multiple Integrals*, 1997.
- [199] S. SÖDERHOLM, J. LAHTINEN, C. H. WOLTERS, AND S. PURSIAINEN, *The effects of peeling on finite element method -based EEG source reconstruction*, *Biomedical Signal Processing and Control*, 89 (2024), p. 105695.
- [200] F. TADEL, S. BAILLET, J. C. MOSHER, D. PANTAZIS, AND R. M. LEAHY, *Brainstorm: A User-Friendly Application for MEG/EEG Analysis*, *Computational Intelligence and Neuroscience*, 2011 (2011), p. 879716.
- [201] M. THEVENET, O. BERTRAND, F. PERRIN, T. DUMONT, AND J. PERNIER, *The finite element method for a realistic head model of electrical brain activities: preliminary results*, *Clinical Physics and Physiological Measurement*, 12 (1991), p. 89.

- [202] A. THIELSCHER, A. OPITZ, AND M. WINDHOFF, *Impact of the gyral geometry on the electric field induced by transcranial magnetic stimulation*, *NeuroImage*, 54 (2011), pp. 234–243.
- [203] L. N. TREFETHEN AND D. BAU, *Numerical Linear Algebra, Twenty-fifth Anniversary Edition*, Society for Industrial and Applied Mathematics, Philadelphia, PA, 2022.
- [204] D. S. TUCH, V. J. WEDEEN, A. M. DALE, J. S. GEORGE, AND J. W. BELLIVEAU, *Conductivity tensor mapping of the human brain using diffusion tensor MRI*, *Proceedings of the National Academy of Sciences*, 98 (2001), pp. 11697–11701.
- [205] S. VALLAGHÉ, T. PAPADOPOULOU, AND M. CLERC, *The adjoint method for general EEG and MEG sensor-based lead field equations*, *Physics in Medicine & Biology*, 54 (2008), p. 135.
- [206] A. VALLI, *Solving an electrostatics-like problem with a current dipole source by means of the duality method*, *Applied Mathematics Letters*, 25 (2012), pp. 1410–1414.
- [207] J. S. VAN CAMPEN, E. L. HOMPE, F. E. JANSEN, D. N. VELIS, W. M. OTTE, F. VAN DE BERG, K. P. J. BRAUN, G. H. VISSER, J. W. SANDER, M. JOELS, AND M. ZIJLMANS, *Cortisol fluctuations relate to interictal epileptiform discharges in stress sensitive epilepsy*, *Brain*, 139 (2016), pp. 1673–1679.
- [208] S. VAN DER WALT, J. L. SCHÖNBERGER, J. NÚÑEZ-IGLESIAS, F. BOULOGNE, J. D. WARNER, N. YAGER, E. GOUILLART, T. YU, AND THE SCIKIT-IMAGE CONTRIBUTORS, *scikit-image: image processing in Python*, *PeerJ*, (2014).
- [209] W. VAN DRONGELEN, M. YUCHTMAN, B. D. VAN VEEN, AND A. C. VAN HUFFELEN, *A spatial filtering technique to detect and localize multiple sources in the brain*, *Brain Topography*, 9 (1996), pp. 39–49.
- [210] B. VAN VEEN, W. VAN DRONGELEN, M. YUCHTMAN, AND A. SUZUKI, *Localization of brain electrical activity via linearly constrained minimum variance spatial filtering*, *IEEE Transactions on Biomedical Engineering*, 44 (1997), pp. 867–880.
- [211] E. T. VARGA, D. TERNEY, M. D. ATKINS, M. NIKANOROVA, D. S. JEPPESEN, P. ULDALL, H. HJALGRIM, AND S. BENICZKY, *Transcranial direct current stimulation in refractory continuous spikes and waves during slow sleep: A controlled study*, *Epilepsy Research*, 97 (2011), pp. 142–145.
- [212] F. VATTA, F. MENEGHINI, F. ESPOSITO, S. MININEL, AND F. DI SALLE, *Realistic and Spherical Head Modeling for EEG Forward Problem Solution: A Comparative Cortex-Based Analysis*, *Computational Intelligence and Neuroscience*, 2010 (2010), p. 972060.
- [213] V. VLADIMIROV, *Equations of Mathematical Physics*, *Mir*, 1984.
- [214] M. VÖRÖSLAKOS, Y. TAKEUCHI, K. BRINYICZKI, T. ZOMBORI, A. OLIVA, A. FERNÁNDEZ-RUIZ, G. KOZÁK, Z. T. KINCSES, B. IVÁNYI, G. BUZSÁKI, AND A. BERÉNYI, *Direct effects of transcranial electric stimulation on brain circuits in rats and humans*, *Nature Communications*, 9 (2018), p. 483.
- [215] J. VORWERK, U. AYDIN, C. H. WOLTERS, AND C. R. BUTSON, *Influence of Head Tissue Conductivity Uncertainties on EEG Dipole Reconstruction*, *Frontiers in Neuroscience*, Volume 13 - 2019 (2019).

- [216] J. VORWERK, J.-H. CHO, S. RAMPP, H. HAMER, T. R. KNÖSCHE, AND C. H. WOLTERS, *A guideline for head volume conductor modeling in EEG and MEG*, NeuroImage, 100 (2014), pp. 590–607.
- [217] J. VORWERK, M. CLERC, M. BURGER, AND C. H. WOLTERS, *Comparison of Boundary Element and Finite Element Approaches to the EEG Forward Problem*, Biomedical Engineering / Biomedizinische Technik, 57 (2012), pp. 795–798.
- [218] J. VORWERK, A. HANRATH, C. H. WOLTERS, AND L. GRASEDYCK, *The multipole approach for EEG forward modeling using the finite element method*, NeuroImage, 201 (2019), p. 116039.
- [219] J. VORWERK, R. OOSTENVELD, M. C. PIASTRA, L. MAGYARI, AND C. H. WOLTERS, *The FieldTrip-SimBio pipeline for EEG forward solutions*, BioMedical Engineering On-Line, 17 (2018), p. 37.
- [220] J. VORWERK, C. H. WOLTERS, AND D. BAUMGARTEN, *Global sensitivity of EEG source analysis to tissue conductivity uncertainties*, Frontiers in Human Neuroscience, Volume 18 - 2024 (2024).
- [221] M. VOSS, R. ASENJO, AND J. REINDERS, *Pro TBB: C++ Parallel Programming with Threading Building Blocks*, 01 2019.
- [222] J. VRBA AND S. ROBINSON, *Linearly constrained minimum variance beamformers, synthetic aperture magnetometry, and MUSIC in MEG applications*, in Conference Record of the Thirty-Fourth Asilomar Conference on Signals, Systems and Computers (Cat. No.00CH37154), vol. 1, 2000, pp. 313–317 vol.1.
- [223] M. WAGNER, M. FUCHS, AND J. KASTNER, *SWARM: sLORETA-weighted accurate minimum norm inverse solutions*, International Congress Series, 1300 (2007), pp. 185–188. New Frontiers in Biomagnetism. Proceedings of the 15th International Conference on Biomagnetism, Vancouver, BC, Canada, August 21-25, 2006.
- [224] S. WAGNER, M. BURGER, AND C. H. WOLTERS, *An Optimization Approach for Well-Targeted Transcranial Direct Current Stimulation*, SIAM Journal on Applied Mathematics, 76 (2016), pp. 2154–2174.
- [225] S. WAGNER, F. LUCKA, J. VORWERK, C. HERRMANN, G. NOLTE, M. BURGER, AND C. WOLTERS, *Using reciprocity for relating the simulation of transcranial current stimulation to the EEG forward problem*, NeuroImage, 140 (2016), pp. 163–173. Transcranial electric stimulation (tES) and Neuroimaging.
- [226] W. WANG AND S. EISENBERG, *A three-dimensional finite element method for computing magnetically induced currents in tissues*, IEEE Transactions on Magnetics, 30 (1994), pp. 5015–5023.
- [227] D. WEINSTEIN, L. ZHUKOV, AND C. JOHNSON, *Lead-field Bases for Electroencephalography Source Imaging*, Annals of Biomedical Engineering, 28 (2000), pp. 1059–1065.
- [228] J. WELLMER, Y. PARPALEY, S. RAMPP, S. POPKIROV, H. KUGEL, ÜMIT AYDIN, C. H. WOLTERS, M. VON LEHE, AND J. VOGES, *Lesion guided stereotactic radiofrequency thermocoagulation for palliative, in selected cases curative epilepsy surgery*, Epilepsy Research, 121 (2016), pp. 39–46.

- [229] K. J. WERHAHN, E. HARTL, K. HAMANN, M. BREIMHORST, AND S. NOACHTAR, *Latency of interictal epileptiform discharges in long-term EEG recordings in epilepsy patients*, Seizure, 29 (2015), pp. 20–25.
- [230] S. WEST, S. J. NOLAN, AND R. NEWTON, *Surgery for epilepsy: a systematic review of current evidence*, Epileptic Disorders, 18 (2016), pp. 113–121.
- [231] B. U. WESTNER, S. S. DALAL, A. GRAMFORT, V. LITVAK, J. C. MOSHER, R. OOSTENVELD, AND J.-M. SCHOFFELEEN, *A unified view on beamformers for M/EEG source reconstruction*, NeuroImage, 246 (2022), p. 118789.
- [232] S. WIEBE AND N. JETTE, *Pharmacoresistance and the role of surgery in difficult to treat epilepsy*, Nature Reviews Neurology, 8 (2012), pp. 669–677.
- [233] H. WILSON, A. MOISEEV, S. PODIN, AND M. QURAAN, *Continuous head localization and data correction in MEG*, International Congress Series, 1300 (2007), pp. 623–626. New Frontiers in Biomagnetism. Proceedings of the 15th International Conference on Biomagnetism, Vancouver, BC, Canada, August 21-25, 2006.
- [234] S. B. WILSON AND R. EMERSON, *Spike detection: a review and comparison of algorithms*, Clinical Neurophysiology, 113 (2002), pp. 1873–1881.
- [235] D. WILTON, S. RAO, A. GLISSON, D. SCHAUBERT, O. AL-BUNDAK, AND C. BUTLER, *Potential integrals for uniform and linear source distributions on polygonal and polyhedral domains*, IEEE Transactions on Antennas and Propagation, 32 (1984), pp. 276–281.
- [236] B. WIRTH, *Green’s Function for the Neumann–Poisson Problem on n -Dimensional Balls*, The American Mathematical Monthly, 127 (2020), pp. pp. 737–743.
- [237] C. WOLTERS, A. ANWANDER, X. TRICOCHÉ, D. WEINSTEIN, M. KOCH, AND R. MACLEOD, *Influence of tissue conductivity anisotropy on EEG/MEG field and return current computation in a realistic head model: A simulation and visualization study using high-resolution finite element modeling*, NeuroImage, 30 (2006), pp. 813–826.
- [238] C. WOLTERS, L. GRASEDYCK, AND W. HACKBUSCH, *Efficient computation of lead field bases and influence matrix for the FEM-based EEG and MEG inverse problem*, Inverse Problems, 20 (2004), pp. 1099–1116.
- [239] C. WOLTERS, H. KÖSTLER, C. MÖLLER, J. HÄRDTLEIN, L. GRASEDYCK, AND W. HACKBUSCH, *Numerical mathematics of the subtraction method for the modeling of a current dipole in EEG source reconstruction using finite element head models*, SIAM J. Scientific Computing, 30 (2007), pp. 24–45.
- [240] A. WOODS, A. ANTAL, M. BIKSON, P. BOGGIO, A. BRUNONI, P. CELNIK, L. COHEN, F. FREGNI, C. HERRMANN, E. KAPPENMAN, H. KNOTKOVA, D. LIEBETANZ, C. MINIUSSI, P. MIRANDA, W. PAULUS, A. PRIORI, D. REATO, C. STAGG, N. WENDEROTH, AND M. NITSCHKE, *A technical guide to tDCS, and related non-invasive brain stimulation tools*, Clinical Neurophysiology, 127 (2016), pp. 1031–1048.
- [241] D. YANG, Q. WANG, C. XU, F. FANG, J. FAN, L. LI, Q. DU, R. ZHANG, Y. WANG, Y. LIN, Z. HUANG, H. WANG, C. CHEN, Q. XU, Y. WANG, Y. ZHANG, Z. ZHANG, X. ZHAO, X. ZHAO, T. LI, C. LIU, Y. NIU, Q. ZHOU, Q. ZHOU, Y. DUAN, X. LIU, T. YU, Q. XUE, J. LI, X. DAI, J. HAN, C. REN, H. XU, N. LI, J. ZHANG, N. XU, K. YANG, AND Y. WANG, *Transcranial direct current stimulation reduces seizure frequency*

- in patients with refractory focal epilepsy: A randomized, double-blind, sham-controlled, and three-arm parallel multicenter study*, Brain Stimulation, 13 (2020), pp. 109–116.
- [242] R. YTTERLID AND E. SHELLSHEAR, *BVH Split Strategies for Fast Distance Queries*, Journal of Computer Graphics Techniques (JCGT), 4 (2015), pp. 1–25.

AKADEMISCHER LEBENSLAUF

Name:	Malte Bernhard Höltershinken
geboren am:	13.01.1997 in Lingen
Staatsangehörigkeit:	deutsch
Hochschulreife (Abitur):	am 30.06.2015 in Meppen
Studium:	Mathematik 01.10.2015 bis 31.03.2022 an der Universität Münster
Abschlüsse:	Bachelor of Science: 19.09.2018 Master of Science: 27.09.2021
Promotion:	seit 01.10.2021 am Institut für Angewandte Mathematik, bei Prof. Dr. Christian Engwer und Prof. Dr. Carsten H. Wolters (vom Fachbereich Medizin)
Beschäftigung:	seit 01.10.2021 als wissenschaftlicher Mitarbeiter am Institut für Biomagnetismus und Biosignalanalyse, Universitätsklinikum Münster