Helsinki University of Technology Institute of Mathematics Research Reports

Espoo 2008 A560

COMPUTATIONAL METHODS IN ELECTROMAGNETIC BIOMEDICAL INVERSE PROBLEMS

Sampsa Pursiainen

TEKNILLINEN KORKEAKOULU
TEKNISKA HÖGSKOLAN
HELSINKI UNIVERSITY OF TECHNOLOGY
TECHNISCHE UNIVERSITÄT HELSINKI
UNIVERSITE DE TECHNOLOGIE D'HELSINKI

Helsinki University of Technology Institute of Mathematics Research Reports
Espoo 2008 A560

COMPUTATIONAL METHODS IN ELECTROMAGNETIC BIOMEDICAL INVERSE PROBLEMS

Sampsa Pursiainen

Dissertation for the degree of Doctor of Science in Technology to be presented, with due permission of the Faculty of Information and Natural Sciences, for public examination and debate in auditorium K at Helsinki University of Technology (Espoo, Finland) on the 18th of December 2008, at 12 o'clock noon.

Helsinki University of Technology
Faculty of Information and Natural Sciences
Department of Mathematics and Systems Analysis

Sampsa Pursiainen
Department of Mathematics and Systems Analysis
Helsinki University of Technology
P.O. Box 1100, FI-02015 TKK, Finland
E-mail: sampsa.pursiainen@tkk.fi

ISBN 978-951-22-9679-8 (print)
ISBN 978-951-22-9680-4 (PDF)
ISSN 0784-3143 (print)
ISSN 1797-5867 (PDF)
TKK Mathematics, 2008

Helsinki University of Technology
Faculty of Information and Natural Sciences
Department of Mathematics and Systems Analysis
P.O. Box 1100, FI-02015 TKK, Finland
email: math@tkk.fi http://math.tkk.fi/

Sampsa Pursiainen: *Computational methods in electromagnetic biomedical inverse problems*; Helsinki University of Technology Institute of Mathematics Research Reports A560 (2008).

Abstract: *This work concerns computational methods in electromagnetic biomedical inverse problems. The following biomedical imaging modalities are studied: electro/magnetoencephalography (EEG/MEG), electrical impedance tomography (EIT), and limited-angle computerized tomography (limited-angle CT). The use of a priori information about the unknown feature is necessary for finding an adequate answer to an inverse problem. Both classical regularization techniques and Bayesian methodology are applied to utilize the a priori knowledge. The inverse problems specifically considered in this work include determination of relatively small electric conductivity anomalies in EIT, dipole-like sources in EEG/MEG, and multiscale X-ray absorbing structures in limited-angle CT. Computational methods and techniques applied for solving these have been designed case-by-case. Results concern, among others, appropriate parametrization of inverse problems; two-stage reconstruction processes, in which a region of interest (ROI) is determined in the first stage and the actual reconstruction is found in the second stage; effective forward simulation through h- and p- versions of the finite element method (FEM); localization of dipole-like electric sources through hierarchical Bayesian models; implementation of the Kirsch factorization method for reconstruction of conductivity anomalies; as well as the use of a coarse-to-fine reconstruction strategy in linear inverse problems.*

AMS subject classifications: Primary 92C55; Secondary 78A70, 15A29, 35R30, 65R32

Keywords: inverse problems, electromagnetics, biomedical imaging, electroencephalography/magnetoencephalography (EEG/MEG), electrical impedance tomography (EIT), limited-angle computerized tomography (limited-angle CT), regularization, Bayesian methodology, Markov chain Monte Carlo (MCMC), factorization method of Kirsch, forward modeling and simulation, h-FEM, p-FEM

Sampsa Pursiainen: *Laskennallisia menetelmiä biolääketieteellisissä sähkömagneettisissa inversio-ongelmissa*

Tiivistelmä: *Tässä työssä käsitellään biolääketieteen sähkömagneettisia käänteisongelmia. Sovelluskohteina ovat seuraavat biolääketieteen kuvantamismenetelmät: elektro-/magnetoenkefalografia (EEG/MEG), impedanssitomografia (EIT) ja rajoitetun kulman tietokonetomografia (rajoitetun kulman CT). Käänteisongelman ratkaisemisessa on välttämätöntä käyttää a priori -tietoa kohteena olevasta tuntemattomasta. A priori -tiedon hyöntämiseen käytetään tässä työssä sekä perinteisiä regularisointitekniikoita että Bayeslaisia menetelmiä. Tässä työssä erityisesti käsiteltävät inversio-ongelmat ovat pienten sähköisen johtavuuden poikkeamien (EIT), dipolin kaltaisten sähköisten lähteiden (EEG/MEG) ja karkeiden röntgensäteitä absorboivien rakenteiden (rajoitetun kulman CT) havaitseminen. Sovelletut laskennalliset menetelmät ja tekniikat ovat tapauskohtaisesti suunniteltuja. Tulokset käsittelevät mm. käänteisongelmien parametrisointiin sopivia tapoja, kaksivaiheista rekonstruointia määrittämällä kiinnostuksen kohteena oleva joukko (ROI) ensimmäisessä vaiheessa ja varsinainen rekonstruktio toisessa vaiheessa, suoran ongelman simulointia elementtimenetelmän (FEM) h- ja p-versioiden avulla, dipolin kaltaisten sähkökentän lähteiden paikantamista Bayeslaisilla menetelmillä, Kirschin faktorointimenetelmän soveltamista johtavuus-anomalioiden rekonstruktiointiin sekä asteittain tarkennettavan rekonstruktion laskemista lineaarisissa käänteisongelmissa.*

Avainsanat: käänteisongelmat, sähkömagnetiikka, biolääketieteellinen kuvantaminen, elektroenkefalografia/magnetoenkefalografia (EEG/MEG), impedanssitomografia (EIT), rajoitetun kulman tietokonetomografia (rajoitetun kulman CT), regularisointi, Bayeslaiset menetelmät, Markov-ketju Monte Carlo (MCMC), Kirschin faktorointimenetelmä, suoran ongelman mallinnus ja simulaatio, h-FEM, p-FEM

"I think it is good that books still exist, but they make me sleepy."
— Frank Zappa (1940–1993), quoted from F Zappa and P Occhiogrosso.
The Real Frank Zappa Book. Poseidon Press, New York, 1989.

Contents

Papers

[**I**] S Pursiainen. Two-stage reconstruction of a circular anomaly in electrical impedance tomography. *Inverse Problems*, 22(5):1689–1703, 2006.

[**II**] S Pursiainen and H Hakula. A High-order Finite Element Method for Electrical Impedance Tomography. *PIERS Online*, 2(3):260–264, 2006.

[**III**] N Hyvönen, H Hakula, and S Pursiainen. Numerical implementation of the factorization method within the complete electrode model of electrical impedance tomography. *Inverse Problems and Imaging*, 1(2):299–317, 2007.

[**IV**] S Pursiainen. EEG/MEG forward simulation through h- and p-type finite elements. *Journal of Physics: Conference Series*, 124(1), 012041 (11pp), 2008.

[**V**] D Calvetti, H Hakula, S Pursiainen, and E Somersalo. Conditionally Gaussian hypermodels for source localization in brain imaging. arXiv:0811.3185.

[**VI**] S Pursiainen. Coarse-to-fine reconstruction in linear inverse problems with application to limited-angle computerized tomography. *Journal of Inverse and Ill-Posed Problems*, To appear.

Own contribution to papers

[**I**] This paper is an independent research of the author.

[**II**] The idea of the article is based on [**I**]. The article is written by SP, while the numerical results were obtained using a computer code made by HH.

[**III**] The main contribution of SP is the development of the forward simulation reported in [**II**]. Part of the paper is written by SP.

[**IV**] This paper reports independent research of the author.

[**V**] The contribution of SP is the numerical experiments and results concerning the realistic geometry. The realistic head model and the forward simulation applied in this study were developed during the study of [**IV**]. The adaptation of the inversion algorithm to include the realistic geometry is due to SP. Part of the paper is written by SP.

[**VI**] This paper reports independent research of the author.

List of abbreviations

The abbreviations used include the following:

ARD	Automatic relevance determination
BEM	Boundary element method
CEM	Complete electrode model
CG	Conjugate gradient
CM	Conditional mean
CM	Continuum model (in Paper [**III**])
CPU	Central processing unit
CSF	Cerebrospinal fluid
CT	Computerized tomography
DOT	Diffuse optical tomography
ECG	Electrocardiography
EEG	Electroencephalography
EIT	Electrical impedance tomography
EM	Expectation maximization
FBP	Filtered back projection
FEM	Finite element method
FGMRES	Flexible generalized minimal residual
FOCUSS	Focal underdetermined system solver
h-**FEM**	Finite element method (h-version)
IAS	Iterative alternating sequential
LORETA	Low resolution electromagnetic tomography
MAP	Maximum *a posteriori*
MCE	Minimum current estimate
MCG	Magnetocardiography
MCMC	Markov chain Monte Carlo
MEG	Magnetoencephalography
MNE	Minimum norm estimate
MRI	Magnetic resonance imaging
MSE	Minimum support estimate
p-**FEM**	Finite element method (p-version)
PCG	Preconditioned conjugate gradient
PSNR	Peak signal-to-noise ratio
ROI	Region of interest
SVD	Singular value decomposition

Acknowledgements

This work covers nearly five years of research with the principal goal to discover novel aspects of computational methods and techniques in electromagnetic biomedical inverse problems. This is admittedly an ambitious aim, since in science and life in general, true novelty and value of ideas can often be evaluated only afterwards. In terms of science, proper evaluation of a work always requires knowledge that is rather of *a posteriori* than of *a priori* nature. Working on these studies has been enjoyable and I have also felt privileged to learn new things about this particular field of science.

I want to express my gratitude for my supervisor Professor Erkki Somersalo, PhD, Case Western Reserve University, Cleveland, Ohio. Also Lecturer Dr. Harri Hakula, DTech, Helsinki University of Technology, had a central role in the studies as a co-author and a close advisor for which I wish to thank him. Further thanks go to the other co-authors Professor Daniela Calvetti, PhD, Case Western Reserve University, Cleveland, Ohio, and Dr. Nuutti Hyvönen, DTech, Helsinki University of Technology, for their valuable work. I also want to thank Dr Carsten Wolters, PhD, University of Münster, and Dr. Jari Toivanen, PhD, Standford University, Stanford, California, for pre-examining this work; Professor Michele Piana, PhD, University of Verona and University of Genoa, for accepting the invitation to serve as an opponent; Professor Olavi Nevanlinna, DTech, Helsinki University of Technology, Director of the Institute of Mathematics, and the Administration of the Institute Mathematics for giving me the opportunity to grow as a research scientist at the Institute; as well as all the colleagues and friends for valuable discussions and advice. The Graduate School of Applied Electromagnetism (Ministry of Education, Finland, project number 77014) is thanked for financial support.

Part I

Overview

1 Introduction

This work concerns computational methods in electromagnetic biomedical inverse problems. The term inverse problem [29] refers to a typically ill-posed and ill-conditioned problem to estimate or reconstruct an unknown feature (parameter) based on the available data. Generally, any question difficult to answer to the full satisfaction can be considered as an inverse problem if the process generating the data can be modeled mathematically. Inverse problems have applications in many fields of science, particularly in physics but increasingly also in other fields from biology to linguistics. Electromagnetic biomedical inverse problems belong to the field of biomedical engineering [17] the goal being to retrieve information by non-invasive electromagnetic measurements. These arise, for instance, in electro/magnetoencephalography (EEG/MEG) [24], electrical impedance tomography (EIT) [11, 13], and limited-angle computerized tomography (limited-angle CT) [14, 22], which constitute the biomedical imaging modalities studied in this work.

EEG/MEG, EIT, and limited-angle CT correspond to inverse problems of source determination, sounding, and high-frequency imaging, respectively [29]. In inverse source problems, the primary source of an electromagnetic field is to be estimated based on measurements of some components of the field outside the source domain. In EEG/MEG, the primary current density generated by neural activity in the brain is to be reconstructed from external potential/magnetic field measurements. In sounding problems, the electromagnetic field is generated artificially for measurement purposes and, based on the measurements, some electromagnetic properties of the interacting target are to be recovered. This is the situation in EIT, in which a conductivity distribution in an object is to be reconstructed from voltage measurement data on the boundary. Again, in high-frequency imaging, an object is exposed to high-frequency electromagnetic radiation, like X-rays in CT, and the problem is to estimate the absorption properties of the target based on the transmission data. In limited-angle CT, the estimate is computed from a number of X-ray projection images with a limited range of directions of illumination. In the present versions of EEG/MEG and limited-angle CT, the inverse problems are linear, meaning that the dependence of the data on the parameter to be estimated can be written as a linear system. The EIT inverse problem is, in contrast, non-linear.

Due to the ill-posed and ill-conditioned nature of inverse problems a unique solution seldom exists and erroneous estimates can result due to errors in data or small uncertainties in the forward model. Consequently, to find an adequate answer to an inverse problem the use of *a priori* information about the unknown feature is necessary. Two popular approaches to utilize *a priori* knowledge are classical regularization techniques [29] and Bayesian methodology [29, 38], both of which are applied in this work. Although there are some similarities between these approaches, they are philosophically different. Whereas a classical regularization technique is aimed at producing a single reasonable estimate for the unknown feature, the goal in the Bayesian methodology is to define one or more subjective posterior probability distributions for the unknown and to make statistical inference based on the posterior distributions. The posterior distribution is often used to derive pointwise estimates, such as maximum a posteriori (MAP) or conditional mean (CM) estimates. One of the advantages in the Bayesian approach is that the *a priori* information is incorporated into the prior density, whose product with the likelihood of the observed data is proportional to the posterior.

1.1 Aims and scope

The aim of this work is to discover novel aspects of both inverse and forward computations to effectively solve electromagnetic biomedical inverse problems. Due to the diversity of problem types, several applications were studied. The inverse problems specifically considered include the determination of relatively small electric conductivity anomalies in EIT, focal electric sources in EEG/MEG, and multiscale X-ray absorbing structures in limited-angle CT. Appropriate computational methods and techniques applied for solving these were designed case-by-case.

The applied methods depend on whether the regularization or the Bayesian approach is adopted. This work utilizes Markov chain Monte Carlo (MCMC) integration [33], Tikhonov regularization [29] and preconditioning [42, 21], wavelet filtering [34], quasi-Newton (Gauss-Newton) optimization [29], the conjugate gradient (CG) method [19, 21], the focal underdetermined system solver (FOCUSS) [20, 43], as well as the factorization method [30]. More methods developed for similar purposes can be found, for instance, in [2, 3, 5, 7, 8, 28, 31, 39, 41, 43]. Together with computational inversion methods, this work discusses effective techniques for EIT and EEG/MEG forward simulation [6, 16, 27, 36, 48, 51, 54, 53, 55] through the application of both low-order and high-order versions of the finite element method (h- and p-FEM) [1, 9, 35, 44, 50] to the complete electrode model (CEM) [49]. Forward simulation, in general, is an important part of the numerical solution of inverse problems and is closely related to the researcher's subjective view of the process modeling of the data.

1.2 Structure and contents

This dissertation consists of two parts: (I) an overview covering an introduction to the work, review of the literature, main results, discussion and conclusions, as well as (II) six original papers [**I**]–[**VI**]. The mathematical background of both regularization and Bayesian methodology as well as of the computational methods and electromagnetic applications are briefly reviewed in the overview.

Paper [**I**] considers two-stage reconstruction of a circular anomaly in EIT from the Bayesian point of view. In the two-stage reconstruction process, a region of interest (ROI) containing an anomaly is determined in the first stage, and the actual reconstruction is found in the second stage by exploring the ROI. It is shown that this approach yields good results, when the parametrization of the unknown conductivity distribution is appropriate and MCMC integration based CM estimation together with h-FEM type forward simulation of the CEM are used. Reconstructions obtained this way are compared against reconstructions yielded by quasi-Newton MAP estimation.

EIT forward simulation is studied further in [**II**] which is devoted for h- and p-FEM simulations of the CEM. Paper [**II**] suggests that the p-version of the FEM provides a more effective tool for EIT forward simulation than the standard h-version. Numerical results supporting this view are presented. Forward simulation through high-order finite elements is mentioned in [**I**] as a topic for future work due to problems in h-FEM forward simulation that occurred during the course of that study. In this sense, paper [**II**] is a continuation for [**I**].

The forward simulation developed for [**II**] is successfully applied in [**III**], in which a reconstruction algorithm based on the Kirsch factorization method is implemented and tested. Paper [**III**] shows that the factorization method can be applied for EIT and its functionality is demonstrated through results of numerical experiments, in which both circular and kite-shaped anomalies were successfully reconstructed.

Paper [**IV**] studies h- and p-FEM type EEG/MEG forward simulations, both based on the CEM. A three-dimensional geometrically realistic finite element mesh, specifically designed for this study, is utilized. Performances are compared through direct measurement of the forward simulation error as well as through reconstructions of multiple dipole-like sources computed using a regularized FOCUSS algorithm. It is shown that in terms of the relative discretization error the performance of the p-version FEM is superior to that of the h-version.

Paper [**V**] utilizes the forward simulation and the realistic finite element mesh implemented during the study of [**IV**]. The focus of [**V**] is on Bayesian hypermodels for source localization applied to EEG/MEG; a generalized gamma distribution is used as a hyperprior and computation of MAP and CM estimates is considered. An iterative alternating sequential (IAS) algorithm for MAP estimation is proposed and tested. It is shown that for different choices of parameters, the IAS algorithm coincides or almost coincides with a number of widely used estimation strategies, e.g. the FOCUSS. MAP estimates produced via the IAS are compared against MCMC based CM estimates in reconstruction of multiple dipole-like sources. The results suggest, in general, that CM estimation is superior over MAP estimation for measuring activity in different parts of the brain.

Finally, in [**VI**], a coarse-to-fine reconstruction procedure for linear inverse problems is proposed and studied. This procedure is partially based on iterative regularization through CG optimization. A nearly block diagonal representation of a quadratic form to be optimized is obtained via wavelet filtering and Tikhonov preconditioning. In the numerical experiments, the proposed procedure is applied to limited-angle CT. The results show that a combined use of wavelet filters and preconditioning can be effective within the studied class of inverse problems.

2 Literature review

2.1 Forward model and simulation

Existence of a mathematical forward model is necessary in both regularization and Bayesian approaches to inverse problems. It is assumed in this work that one can define a deterministic or probabilistic forward mapping [29] of the form

$$\mathbf{y} = \mathbf{F}(\mathbf{x}, \mathbf{z}, \mathbf{n}), \tag{2.1}$$

in which $\mathbf{y}$ denotes the data; $\mathbf{x}, \mathbf{z}$ denote parameters corresponding to unknown interesting and uninteresting features affecting the data outcome, respectively, and $\mathbf{n}$ describes the noise contamination of the data. It is also assumed that recovery of $\mathbf{x}$, given $\mathbf{y}$, is an ill-posed and ill-conditioned problem independently of whether $\mathbf{z}$ and $\mathbf{n}$ are given or not. This means that $\mathbf{x}$ is not uniquely solvable as well as that a minor uncertainty on $\mathbf{F}$, $\mathbf{z}$, or $\mathbf{n}$ can lead to a major uncertainty on $\mathbf{x}$. For these reasons, it is computationally infeasible, for instance, to directly solve (2.1) or to search the classical least-squares (LS) solution given by $\hat{\mathbf{x}} = \arg\min_{\mathbf{x}} \|\mathbf{y} - \mathbf{F}(\mathbf{x}, \mathbf{z}, \mathbf{n})\|_2^2$.

2.1.1 Interpretation and construction

A mapping of the form (2.1) can be interpreted to represent a forward model, or a forward simulation, depending on whether an abstract mathematical model or a numerical simulation is in question. This work considers mainly the latter interpretation, and therefore, all the variables in (2.1) are assumed to be real and finite dimensional:

$$\mathbf{x} \in \mathbb{R}^n, \quad \mathbf{y}, \mathbf{n} \in \mathbb{R}^m, \quad \text{and} \quad \mathbf{z} \in \mathbb{R}^M. \tag{2.2}$$

How a forward model or a forward simulation is constructed in practice, is highly dependent on the application. In EEG/MEG and EIT, forward models are usually based on the quasi-static approximation of the Maxwell's equations [24, 26, 29], which can be simulated numerically e.g. through the FEM (Figure 2.1). In contrast, the forward model of the limited-angle CT comprises the Radon transform and the forward simulation is usually based on linear pixel or voxel sums approximating the line integrals of the Radon transform [29].

2.2 Regularization

The goal in regularization is, through the use of *a priori* information, to produce a reasonable estimate of $\mathbf{x}$ that is in good agreement with $\mathbf{y}$ in the sense of (2.1). Regularization is a non-probabilistic way to solve inverse problems and it is, therefore, assumed in this section that the forward mapping does

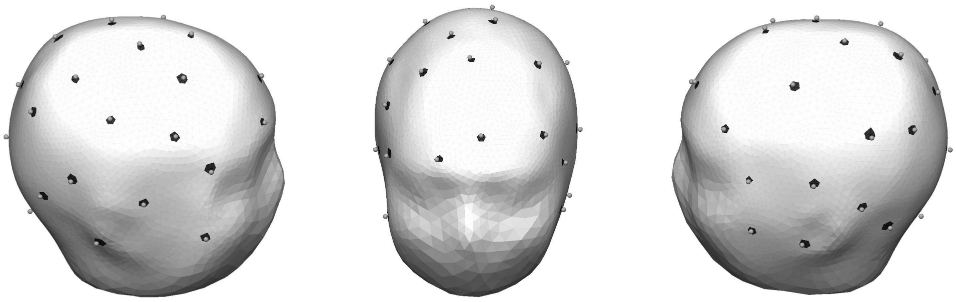

Figure 2.1: The outer surface of a finite element mesh approximating a human head. This mesh can be used in both EIT and EEG/MEG. The darkened surface patches show locations of contact electrodes and the grey spheres over the patches show locations of magnetic field sensors.

not involve any uncertainty, i.e. that it is a deterministic function of the form $\mathbf{y} = \mathbf{F}(\mathbf{x})$.

Classical regularization methods [29] are usually based on (i) use of regularizing functions, (ii) use of iterative regularization, and (iii) projecting the solution space to some subspace. Often two or more of these regularization techniques are used together.

2.2.1 Regularizing functions

Incorporating a regularizing function into the reconstruction process usually means finding the minimizer of

$$\Psi(\mathbf{x} \mid \mathbf{y}) = \|\mathbf{W}^{-1/2}(\mathbf{y} - \mathbf{F}(\mathbf{x}))\|_2^2 + G(\mathbf{x}), \tag{2.3}$$

where $\mathbf{W}^{-1/2}$ is the inverse of the square-root of a symmetric and positive definite weighting matrix, $G(\mathbf{x})$ is a regularizing function and $\mathbf{y}$ is the data, possibly corrupted by errors. In this regularization scheme, it is important that the regularizing function is chosen so that $\mathbf{x} \to \Psi(\mathbf{x} \mid \mathbf{y})$ has a unique minimizer and that a reasonable balance between the quality of the estimates and the numerical stability of the minimization process is achieved.

The minimization of (2.3) is called generalized Tikhonov regularization, if the regularizing function is of the form $G(\mathbf{x}) = \alpha\|\mathbf{G}\mathbf{x}\|_2^2$, where $\alpha > 0$ is a regularization parameter to be chosen appropriately. If $G(\mathbf{x}) = \alpha\|\mathbf{x}\|_2^2$, this approach is often called the classical Tikhonov regularization. In linear inverse problems, where $\mathbf{F}(\mathbf{x}) = \mathbf{F}\mathbf{x}$, Tikhonov regularization is widely used as the corresponding function to be minimized is a quadratic form, i.e.

$$\Psi(\mathbf{x} \mid \mathbf{y}) = \mathbf{x}^T\mathbf{Q}\mathbf{x} - 2\mathbf{x}^T\mathbf{b} + \mathbf{c}^T\mathbf{c} \tag{2.4}$$

with $\mathbf{Q} = \mathbf{F}^T\mathbf{W}^{-1}\mathbf{F} + \mathbf{G}^T\mathbf{G}$, $\mathbf{b} = \mathbf{F}^T\mathbf{W}^{-1}\mathbf{y}$, and $\mathbf{c} = \mathbf{W}^{-1/2}\mathbf{y}$, and the minimizer can be obtained directly by solving the linear system

$$\mathbf{Q}\mathbf{x} = \mathbf{b}, \tag{2.5}$$

whose solution is unique if $\mathbf{G}$ is appropriately chosen. In typical applications, the Tikhonov regularization tends to produce blurred estimates and, therefore, it is often not preferred, when the feature to be estimated is known *a priori* to contain sharp-edged structures.

Sharp structured estimates can be produced, for instance, by choosing $G(\mathbf{x}) = \alpha\|\mathbf{x}\|_1$, which leads to a non-linear minimization problem. This is the choice in the regularized FOCUSS algorithm [20, 43], in which the minimizer is found through a relatively simple iterative relaxation process described by $\mathbf{x}^{(0)} = (1, 1, \ldots, 1)^T$, $\mathbf{W}_k = \text{diag}(|x_1^{(k)}|, |x_2^{(k)}|, \ldots, |x_n^{(k)}|)^{1/2}$, $\mathbf{F}_k = \mathbf{W}^{-1/2}\mathbf{F}\mathbf{W}_k$, and

$$\mathbf{x}^{(k+1)} = \mathbf{W}_k(\mathbf{F}_k^T\mathbf{F}_k + \frac{\alpha}{2}\mathbf{I})^{-1}\mathbf{F}_k^T\mathbf{W}^{-1/2}\mathbf{y}, \quad k = 0, 1, \ldots, n, \tag{2.6}$$

where $n \in \mathbb{N}$ is the user specified number of iterations. The regularized FOCUSS tends to find a sparse estimate, which is focused in terms of non-zero entries. Therefore, it is a suitable reconstruction strategy for finding sparsely distributed and well-localized sources.

When the forward mapping is non-linear, one potentially applicable optimization method is the following quasi-Newton (Gauss-Newton) iteration [29]:

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} - \lambda[D\nabla\Psi(\mathbf{x}^{(k)} \mid \mathbf{y})]^{-1}\nabla\Psi(\mathbf{x}^{(k)} \mid \mathbf{y}),$$

where λ is a step size control parameter to be specified by the user. This iteration can be described as Newton's method applied for finding the zero of $\nabla\Psi(\mathbf{x} \mid \mathbf{y})$. Since quasi-Newton optimization is based on the assumption that $\Psi(\mathbf{x} \mid \mathbf{y})$ is two times differentiable, it is particularly well suited for cases in which the reconstruction needs to be smooth.

Reconstruction via factorization method

The reconstruction algorithm below, suitable for localization of anomalies, is based on the factorization method introduced by Kirsch [30]. It constitutes an example of how regularizing functions can be used in a different way.

Given the data $\mathbf{y}$, the matrix $\mathbf{F}$ corresponding to the linear forward mapping, a discrepancy parameter $\delta > 0$, as well as a thresholding parameter $\beta > 0$; set $k = 1$ and $\mathcal{S} = \emptyset$, and while $k < n$, repeat the following steps:

- Pick the kth standard basis vector $\mathbf{e}_k = (0, \ldots, 0, 1, 0, \ldots, 0)$ and find $\tau > 0$ such that [52] the equations $(\tau\mathbf{I} + (\mathbf{F}^T\mathbf{F})^{1/2})\mathbf{x}_\delta = (\mathbf{F}^T\mathbf{F})^{1/4}\mathbf{e}_k$ and $\|(\mathbf{F}^T\mathbf{F})^{1/4}\mathbf{x}_\delta - \mathbf{e}_k\| = \delta\|\mathbf{x}_\delta\|$ are satisfied with some $\mathbf{x}_\delta$.
- If $\tau < \beta$, substitute $\mathcal{S} \to \{\mathcal{S}, k\}$, otherwise do nothing.
- Substitute $k \to k + 1$.

The resulting $\mathcal{S}$ describes the set of potentially non-zero vector entries and can, for instance, be used as a reconstruction to represent one or more anomalies. Here, the regularizing function is $\delta\|\mathbf{x}_\delta\|$; the larger is δ, the stronger is the effect of the regularization.

2.2.2 Projections in regularization

The use of projections is a part of classical regularization [32, 42] methods. As an example of how projections can be used in linear inverse problems, consider the singular value truncation [29], in which one first computes the singular value decomposition (SVD) [19] given by

$$\mathbf{F} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T, \tag{2.7}$$

where $\mathbf{U}^T\mathbf{U} = \mathbf{I}$, $\mathbf{V}^T\mathbf{V} = \mathbf{I}$, and $\mathbf{\Sigma} = \mathrm{diag}(\sigma_1, \sigma_2, \ldots, \sigma_n)$ with $\sigma_1 \geq \sigma_2 \geq \ldots \sigma_n \geq 0$, and after that replaces the matrix $\mathbf{F}$ with

$$\widetilde{\mathbf{F}} = \mathbf{U}\widetilde{\mathbf{\Sigma}}\mathbf{V}^T, \tag{2.8}$$

in which $\widetilde{\mathbf{\Sigma}} = \mathrm{diag}(\tilde{\sigma}_1, \tilde{\sigma}_2, \ldots, \tilde{\sigma}_n)$ with $\tilde{\sigma}_i = \sigma_i$, if $\sigma_i \geq \varepsilon > 0$, and $\tilde{\sigma}_i = 0$, otherwise, for $i = 1, 2, \ldots, n$. The parameter ε defines the level of truncation. With a large enough ε, the matrix $\widetilde{\mathbf{F}}$ is well-conditioned and the reconstruction can be obtained, for instance, by computing the classical LS solution. Again, a too large ε leads to blurred reconstructions. In practice, an appropriate choice for ε depends both on the singular values of $\mathbf{F}$ and on the noise level.

The singular value truncation process can be interpreted as a projection of the space of possible solutions, formed by the two orthogonal subspaces

$$S_\varepsilon^+ = \{\, \mathbf{x} \,:\, \|\mathbf{F}\mathbf{x}\| > \varepsilon\|\mathbf{x}\| \,\} \cup \{\mathbf{0}\} \quad \text{and} \quad S_\varepsilon^- = \{\, \mathbf{x} \,:\, \|\mathbf{F}\mathbf{x}\| \leq \varepsilon\|\mathbf{x}\| \,\}, \tag{2.9}$$

onto the former one (S_ε^+). With a very small ε, the subspace S_ε^- can be regarded as the numerical null space [32] of $\mathbf{F}$. Namely, in S_ε^-, the product $\mathbf{F}\mathbf{x}$ can be numerically insensitive to variation of $\mathbf{x}$ and, therefore, recovery of the component of the source vector lying in S_ε^- can be practically impossible, even in a noiseless case.

In some applications, approximations of the subspaces (2.9) can be also obtained based on *a priori* information. For example, wavelets that are well-localized in the frequency domain can be applicable for such a purpose, if according to *a priori* knowledge, the subspaces S_ε^+ and S_ε^- contain mainly low- and high-frequencies, respectively.

2.2.3 Iterative regularization

In iterative regularization, typically the minimizer of

$$\Psi(\mathbf{x} \mid \mathbf{y}) = \|\mathbf{W}^{-1/2}(\mathbf{y} - \mathbf{F}(\mathbf{x}))\|_2^2, \tag{2.10}$$

is approximated by using some iteration with a stopping criterion based on the noise level of the data as well as on how ill-conditioned the problem is. When used for regularization purposes, iterations that converge very rapidly can produce erroneous reconstructions, and, consequently, even very simple and slow converging methods can be advantageous [29].

This work considers minimization of (2.10) through iterative regularization in connection with linear inverse problems. When $\mathbf{F}(\mathbf{x})$ is linear, the function $\Psi(\mathbf{x} \mid \mathbf{y})$ is of the quadratic form (2.4) with $\mathbf{Q} = \mathbf{F}^T\mathbf{W}^{-1}\mathbf{F}$, $\mathbf{b} = \mathbf{F}^T\mathbf{W}^{-1}\mathbf{y}$, and $\mathbf{c} = \mathbf{W}^{-1/2}\mathbf{y}$, meaning that iterative methods for solving the linear system (2.5), e.g. CG [9, 19], can be utilized. The CG based regularization method for regularization purposes can be written explicitly as follows: Choose a stopping parameter δ, an initial guess $\mathbf{x}^0$, set $\mathbf{p}^0 = \mathbf{r}^0 = \mathbf{b} - \mathbf{Q}\mathbf{x}^{(0)}$, $k = 0$, and while $\|\mathbf{r^k}\| \geq \delta$ repeat the steps

$$\alpha_k = -(\mathbf{r}^{(k)})^T\mathbf{r}^{(k)}/(\mathbf{p}^{(k)})^T\mathbf{Q}\mathbf{p}^{(k)}, \tag{2.11}$$
$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \alpha_k\mathbf{p}^{(k)}, \tag{2.12}$$
$$\mathbf{r}^{(k+1)} = \mathbf{r}^{(k)} + \alpha_k\mathbf{Q}\mathbf{p}^{(k)}, \tag{2.13}$$
$$\beta_k = (\mathbf{r}^{(k+1)})^T\mathbf{r}^{(k+1)}/(\mathbf{r}^{(k)})^T\mathbf{r}^{(k)}, \tag{2.14}$$
$$\mathbf{p}^{(k+1)} = \mathbf{r}^{(k+1)} + \beta_k\mathbf{p}^{(k)}. \tag{2.15}$$

The reconstruction is the final approximation $\mathbf{x}^{(k+1)}$.

Preconditioning

In linear inverse problems, results obtained using standard iterative regularization methods can often be enhanced through preconditioning [12, 21, 42]. In contrast to the traditional goal of preconditioning, which is to speed up the convergence rate, the objective of preconditioning inverse problems is to reorganize the problem into a better conditioned form in order to obtain better reconstructions or to produce estimates with desired features.

Often, the objective is to choose the weighting matrix $\mathbf{W}$ so that the singular values of $\mathbf{Q}$ are either exactly zero or close to one, which means that $\mathbf{Q}$ almost or exactly satisfies the equation

$$\mathbf{V}^T\mathbf{Q}\mathbf{V} = \begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}, \tag{2.16}$$

where $\mathbf{V}$ is a matrix containing the right singular vectors of $\mathbf{F}$ as in (2.7). This can be explained by examining iterative minimization of a two-dimensional quadratic form with contour ellipses (Figure 2.2) that are very eccentric, i.e. the semimajor axes are much longer than the semiminor axes. Due to the eccentricity, the graph of the quadratic form is flat or almost flat in the direction of the semimajor axes, in which case even a well implemented minimization algorithm can fail to find the minimum, especially if the graph is almost flat in the direction of each basis vector (Figure 2.2). The imbalance between the axes lengths does not exist, at least in theory, if an appropriate preconditioning is applied to scale the axes.

One possible preconditioning method is Tikhonov preconditioning, which is based on the classical Tikhonov regularization. In this preconditioning scheme, the weighting matrix is of the form

$$\mathbf{W} = (\mathbf{F}\mathbf{F}^T + \delta\mathbf{I})^{-1}, \tag{2.17}$$

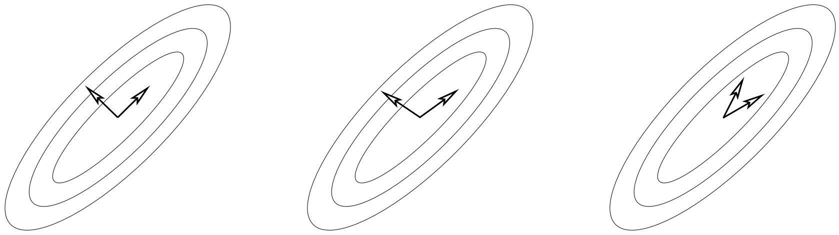

Figure 2.2: Contour ellipses of a two-dimensional quadratic form and three different vector bases in which the quadratic form is diagonal (left), a nearly diagonal (center) and a non-diagonal (right). In the non-diagonal case the graph of the quadratic form can be flat or almost flat in the direction of each basis vector.

where δ is a suitably chosen parameter. It can be shown via the singular value decomposition (SVD) of $\mathbf{F}$ that when δ is small enough, the matrix $\mathbf{Q}$ corresponding to $\mathbf{W}$ almost satisfies (2.16).

2.3 Bayesian inversion

In Bayesian inversion, the goal is to define and study a posterior probability density, as well as to make inferences based on that or on posterior derived estimates. The posterior is even called the Bayesian solution of the specific inverse problem.

Given a prior $p(\mathbf{x}, \mathbf{z}, \mathbf{n})$ containing the *a priori* information; a likelihood function $p(\mathbf{y} \mid \mathbf{x}, \mathbf{z}, \mathbf{n})$, i.e. a conditional density of measuring $\mathbf{y}$; and a marginal likelihood function $p(\mathbf{y}) = \iint p(\mathbf{x}, \mathbf{z}, \mathbf{n})\, p(\mathbf{y} \mid \mathbf{x}, \mathbf{z}, \mathbf{n})\, d\mathbf{n}\, d\mathbf{z}\, d\mathbf{x}$, the posterior density is of the form

$$p(\mathbf{x} \mid \mathbf{y}) = \frac{1}{p(\mathbf{y})} \iint p(\mathbf{x}, \mathbf{z}, \mathbf{n})\, p(\mathbf{y} \mid \mathbf{x}, \mathbf{z}, \mathbf{n})\, d\mathbf{n}\, d\mathbf{z}. \tag{2.18}$$

This is a consequence of the classical Bayes formula $p(a \mid b) = p(a)p(b \mid a)/p(b)$ [29, 38]. If, as is typical, the nuisance parameter $\mathbf{z}$ is not present, the posterior can be written as

$$p(\mathbf{x} \mid \mathbf{y}) = \frac{1}{p(\mathbf{y})} \int p(\mathbf{x}, \mathbf{n})\, p(\mathbf{y} \mid \mathbf{x}, \mathbf{n})\, d\mathbf{n}. \tag{2.19}$$

Note that for a given data $\mathbf{y}$, the product of the prior and the likelihood integrated with respect to $\mathbf{n}$ constitutes the posterior density. Also note that here the data $\mathbf{y}$ is assumed to only affect on the likelihood, not the prior. This is known as the likelihood principle.

2.3.1 Subjective probability and calibration

Probabilities in Bayesian methodology can, in general, be interpreted to be subjective [38]. A subjective probability measures one person's degree of belief, and it can vary from person to person. In Bayesian inversion, this means that the prior and the likelihood can be chosen freely by the researcher according to the best available knowledge on the inverse problem.

For example, if one performs two similar and consecutive experiments producing data vectors $\mathbf{y}_1$ and $\mathbf{y}_2$, respectively, the posterior $p(\mathbf{x} \mid \mathbf{y}_1)$, obtained after the first experiment, can be used as a prior in the second one, which results into a posterior of the form $p(\mathbf{x} \mid \mathbf{y}_1, \mathbf{y}_2)$. This idea of constructing the posterior sequentially in two stages can be effectively applied in Bayesian inversion, since for instance, it is often advantageous to first compute a coarse estimate for $\mathbf{x}$ and then construct the final prior based on that.

This work studies an analogous but not fully Bayesian two-stage approach to posterior construction, which differs from the above described strategy only by that the vectors $\mathbf{y}_1$ and $\mathbf{y}_2$ are assumed to represent the same data. This kind of approach, in which the same data is used twice, can be interpreted as Bayesian bootstrapping; it often works even though it may be in contradiction with the likelihood principle.

2.3.2 Model of independent and additive noise

In the construction of the likelihood function, the role of the applied noise model is central. It is often assumed that the noise $\mathbf{n}$ is independent of $\mathbf{x}$ and $\mathbf{z}$, meaning that the prior is of the form $p(\mathbf{x}, \mathbf{z}, \mathbf{n}) = p(\mathbf{x}, \mathbf{z})\, p_{\text{noise}}(\mathbf{n})$. It is also typical to assume that the noise is additive,

$$\mathbf{y} = \mathbf{F}(\mathbf{x}, \mathbf{z}) + \mathbf{n}, \tag{2.20}$$

where $\mathbf{F}(\mathbf{x}, \mathbf{z})$ is a deterministic function. If both of these assumptions are made, then (2.18) can be rewritten as

$$\begin{aligned} p(\mathbf{x} \mid \mathbf{y}) &= \frac{1}{p(\mathbf{y})} \int p(\mathbf{x}, \mathbf{z}) \int p_{\text{noise}}(\mathbf{n})\, \delta(\mathbf{y} - \mathbf{F}(\mathbf{x}, \mathbf{z}) - \mathbf{n})\, d\mathbf{n}\, d\mathbf{z} && (2.21) \\ &= \frac{1}{p(\mathbf{y})} \int p(\mathbf{x}, \mathbf{z})\, p_{\text{noise}}(\mathbf{y} - \mathbf{F}(\mathbf{x}, \mathbf{z}))\, d\mathbf{z} && (2.22) \end{aligned}$$

where δ denotes the Dirac delta function. Again, if the nuisance parameter $\mathbf{z}$ is not present, then (2.19) is valid and the posterior is given by

$$p(\mathbf{x} \mid \mathbf{y}) = p(\mathbf{x})\, p_{\text{noise}}(\mathbf{y} - \mathbf{F}(\mathbf{x}))/p(\mathbf{y}), \tag{2.23}$$

where $p(\mathbf{x})$ constitutes the prior and $p_{\text{noise}}(\mathbf{y} - \mathbf{F}(\mathbf{x}))$ the likelihood.

In models of independent and additive noise, for example, Poisson and Gaussian densities are popular choices for $p_{\text{noise}}(\mathbf{n})$. This work studies the Gaussian case, that is

$$p_{\text{noise}}(\mathbf{n}) \propto \exp(-\frac{1}{2}\|\mathbf{\Gamma}_{\text{noise}}^{-1/2}(\mathbf{n} - \boldsymbol{\mu}_{\text{noise}})\|_2^2), \tag{2.24}$$

with $\boldsymbol{\mu}_{\text{noise}}$ denoting the mean and $\mathbf{\Gamma}_{\text{noise}}$ a symmetric and positive definite covariance matrix. When $p_{\text{noise}}(\mathbf{n})$ is of the form (2.24) and the posterior is given by (2.23), the likelihood is a Gaussian density with the mean $\mathbf{F}(\mathbf{x}) + \boldsymbol{\mu}_{\text{noise}}$ and covariance matrix $\boldsymbol{\Gamma}_{\text{noise}}$. It is often assumed that $\mathbf{n}$ is Gaussian white noise, meaning that the components are identically distributed zero mean Gaussian random variables, mutually independent.

2.3.3 Priors

In Bayesian inversion, all *a priori* information about $\mathbf{x}$ is assumed to be encoded into the prior density. The choice of the prior has, therefore, an essential influence on the characteristics of the posterior. There exists no simultaneously general and practical strategy to construct a prior. On one hand, incorporating the available prior knowledge into any probability density can be very difficult. On the other hand, there may be several possible alternatives that may perform well. Some extensively used choices are briefly introduced below.

Exponential families

This work considers mainly priors belonging to exponential families, in which all the distributions share a certain exponential form and each individual prior is specified by one or more so-called hyperparameters. Denoting the set of hyperparameters by the vector $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \ldots, \alpha_K)$, a member $p(\mathbf{x}; \boldsymbol{\alpha})$ of an exponential family is given by

$$p(\mathbf{x}; \boldsymbol{\alpha}) = h(\mathbf{x}) \exp\left(\sum_{i=1}^{K} \eta_i(\boldsymbol{\alpha}) T_i(\mathbf{x}) - A(\boldsymbol{\alpha})\right). \tag{2.25}$$

where the functions $h(\mathbf{x}), \eta_i(\boldsymbol{\alpha}), T_i(\mathbf{x})$, and $A(\boldsymbol{\alpha})$ specify the family. Exponential families include e.g. Gaussian, gamma, inverse gamma, Bernoulli, binomial, negative binomial, and Poisson distributions.

A practical example of a prior belonging to an exponential family is a Gaussian smoothness prior,

$$p(\mathbf{x}; \boldsymbol{\mu}, \gamma, \mathbf{L}) \propto \exp(-\frac{1}{2\gamma}\|\mathbf{L}(\mathbf{x} - \boldsymbol{\mu})\|_2^2), \tag{2.26}$$

where the hyperparameter $\boldsymbol{\mu}$ denotes the mean, $\gamma > 0$ controls the variance (peakedness) of the prior, and $\mathbf{L}$ is a symmetric and positive definite matrix that is a discrete approximation of the Laplacian operator. This kind of a prior tends to favor vectors that correspond to spatially smooth structured functions around $\boldsymbol{\mu}$, since for such vectors $\|\mathbf{L}(\mathbf{x} - \boldsymbol{\mu})\|_2^2$ is likely to be small.

A prior of the form (2.25) is closely related to regularizing functions. Namely, denoting the prior by $p(\mathbf{x}) \propto \exp(-G(\mathbf{x}))$ results into a posterior (2.23) of the form

$$p(\mathbf{x} \mid \mathbf{y}) \propto \exp(-\frac{1}{2}\|\boldsymbol{\Gamma}_{\text{noise}}^{-1/2}(\mathbf{y} - \mathbf{F}(\mathbf{x}) - \boldsymbol{\mu}_{\text{noise}})\|_2^2 - G(\mathbf{x})), \tag{2.27}$$

in which the argument of the exponential function is essentially of the form (2.3).

Double exponential priors

Priors that do not belong to an exponential family can produce a posterior of the form (2.27). The example of such a prior is a family of double-exponential priors, the members being

$$p(x; \boldsymbol{\mu}, \kappa) \propto \exp\left(-\kappa \|\boldsymbol{x} - \boldsymbol{\mu}\|_1\right). \tag{2.28}$$

These are fat-tailed densities not forming an exponential family, but which are frequently used as priors when the *a priori* knowledge suggests that $\mathbf{x}$ has only few essentially non-zero entries.

Conjugate priors

Often, it is algebraically advantageous to rely on an existing conjugate prior family. A family $\mathcal{F}$ of prior densities is said to be conjugate to a given likelihood $p(\mathbf{y} \mid \mathbf{x})$, if a posterior distribution of the form $p(\mathbf{x} \mid \mathbf{y}) \propto p(\mathbf{x})p(\mathbf{y} \mid \mathbf{x})$ belongs to $\mathcal{F}$ whenever $p(\mathbf{x})$ belongs to $\mathcal{F}$.

If the likelihood belongs to an exponential family, there exists a conjugate prior which also often is a member of some exponential family. Especially, the family of Gaussian densities is self-conjugate, i.e. Gaussian priors are conjugate to a Gaussian likelihood. For example, Gaussian priors of the form

$$p(\mathbf{x}; \boldsymbol{\mu}_{\text{pr}}, \boldsymbol{\Gamma}_{\text{pr}}) \propto \exp\left(-\frac{1}{2}\|\boldsymbol{\Gamma}_{\text{pr}}^{-1/2}(\mathbf{x} - \boldsymbol{\mu}_{\text{pr}})\|_2^2\right) \tag{2.29}$$

are conjugate to the likelihood function of (2.23) when the noise prior is a Gaussian density of the form (2.24) and $\mathbf{F}(\mathbf{x}) = \mathbf{F}\mathbf{x}$ with some matrix $\mathbf{F}$. The corresponding family of Gaussian posterior densities is specified by the mean and covariance matrix

$$\boldsymbol{\mu}_{\text{post}} = \boldsymbol{\mu}_{\text{pr}} + \boldsymbol{\Gamma}_{\text{pr}}\mathbf{F}^T(\mathbf{F}\boldsymbol{\Gamma}_{\text{pr}}\mathbf{F}^T + \boldsymbol{\Gamma}_{\text{noise}})^{-1}(\mathbf{y} - \mathbf{F}\boldsymbol{\mu}_{\text{pr}} - \boldsymbol{\mu}_{\text{noise}}), \tag{2.30}$$

$$\boldsymbol{\Gamma}_{\text{post}} = \boldsymbol{\Gamma}_{\text{pr}} - \boldsymbol{\Gamma}_{\text{pr}}\mathbf{F}^T(\mathbf{F}\boldsymbol{\Gamma}_{\text{pr}}\mathbf{F}^T + \boldsymbol{\Gamma}_{\text{noise}})^{-1}\mathbf{F}\boldsymbol{\Gamma}_{\text{pr}}. \tag{2.31}$$

That is, both $\boldsymbol{\mu}_{\text{post}}$ and $\boldsymbol{\Gamma}_{\text{post}}$ can be obtained through straightforward linear algebra, which is advantageous with respect to the exploration of the posterior.

Hyperpriors

In some applications, priors can be constructed hierarchically. In a hierarchical model free of nuisance parameters, it is assumed that the parameter of primary interest is of the form $(\mathbf{x}_1, \mathbf{x}_2)$ and the actual prior is a joint density $p(\mathbf{x}_1, \mathbf{x}_2)$ formed by the product of a conditional prior $p(\mathbf{x}_1 \mid \mathbf{x}_2)$ and a hyperprior $p(\mathbf{x}_2)$. The corresponding likelihood $p(\mathbf{y} \mid \mathbf{x}_1, \mathbf{x}_2)$ is hierarchical in the sense that it does not depend on $\mathbf{x}_2$. Therefore, it can be written as $p(\mathbf{y} \mid \mathbf{x}_1)$, and the resulting posterior is given by

$$p(\mathbf{x}_1, \mathbf{x}_2 \mid \mathbf{y}) \propto p(\mathbf{y} \mid \mathbf{x}_1, \mathbf{x}_2)p(\mathbf{x}_1 \mid \mathbf{x}_2)p(\mathbf{x}_2), \tag{2.32}$$

$$= p(\mathbf{y} \mid \mathbf{x}_1)p(\mathbf{x}_1 \mid \mathbf{x}_2)p(\mathbf{x}_2). \tag{2.33}$$

The number of hierarchy levels can be greater than two, since also the hyperprior can be constructed hierarchically.

A hierarchical model can be applicable, for instance, if according to the *a priori* information, the random vector $\mathbf{x}$ is Gaussian with an unknown diagonal covariance matrix. Candidate for the prior could be $p(\mathbf{x}, \gamma) \propto p(\mathbf{x} \mid \gamma)p(\gamma)$ with

$$p(\mathbf{x} \mid \boldsymbol{\gamma}) \propto \exp(-\sum_{\ell=1}^{n} \frac{1}{2\gamma_\ell}(x_\ell - \mu_{\mathrm{pr},\ell})^2), \tag{2.34}$$

$$p(\boldsymbol{\gamma}) = p(\boldsymbol{\gamma}; \alpha, \beta) \propto \exp(-\sum_{\ell=1}^{n} \frac{\beta}{\gamma_\ell} - (\alpha + 1)\sum_{\ell=1}^{n} \log \gamma_\ell). \tag{2.35}$$

Here, the conditional prior of $\mathbf{x}$ is a Gaussian density with the covariance matrix $\boldsymbol{\Gamma}_{\mathrm{pr}} = \mathrm{diag}(\gamma_1, \gamma_2, \ldots, \gamma_n)$, and the variance vector $\boldsymbol{\gamma} = (\gamma_1, \gamma_2, \ldots, \gamma_n)$ is distributed according to the inverse gamma hyperprior specified by the hyperparameters α and β. The family of inverse gamma densities of the form $p(\boldsymbol{\gamma}; \alpha, \beta)$ is conjugate to $p(\mathbf{x} \mid \boldsymbol{\gamma})$, which means that $p(\boldsymbol{\gamma} \mid \mathbf{x})$ is also an inverse gamma density, a fact that can be useful in posterior exploration.

In contrast to the Gaussian smoothness prior (2.26) that favors vectors corresponding to spatially smooth structured functions, the posterior obtained as a product of (2.34) and (2.35) tends to favor vectors having essentially only few non-zero entries, like the double-exponential density (2.28). This is caused by the fat-tailedness of the inverse gamma density, which allows outliers to the variance vector. A posterior with similar characteristics can be obtained, if the inverse gamma hyperprior is replaced with a gamma hyperprior given by

$$p(\boldsymbol{\gamma}; \alpha, \beta) \propto \exp(-\sum_{\ell=1}^{n} \frac{\gamma_\ell}{\beta} + (\alpha - 1)\sum_{\ell=1}^{n} \log \gamma_\ell). \tag{2.36}$$

However, this hyperprior is not conjugate to the conditional prior (2.34).

Weak priors

A prior is called weak, or uninformative, when estimates or inferences based on the posterior are insensitive to perturbations of the prior. In inverse problems, a weak prior leads to a posterior, which is difficult to handle numerically. Therefore, in Bayesian inversion, it is not preferable to apply a weak prior.

For instance, a Gaussian prior with a very large variance is weak. To see this, consider the mean of a Gaussian posterior given by (2.30) with $\boldsymbol{\Gamma}_{\mathrm{pr}} = \gamma_{\mathrm{pr}}\mathbf{I}$, $\boldsymbol{\Gamma}_{\mathrm{noise}} = \gamma_{\mathrm{noise}}\mathbf{I}$, and $\boldsymbol{\mu}_{\mathrm{pr}} = \boldsymbol{\mu}_{\mathrm{post}} = \mathbf{0}$. It holds that

$$\lim_{\gamma_{\mathrm{pr}} \to \infty} \boldsymbol{\mu}_{\mathrm{post}} = \mathbf{F}^{+}\mathbf{y} \quad \text{with} \quad \mathbf{F}^{+} = \lim_{\delta \to 0} \mathbf{F}^T(\mathbf{F}\mathbf{F}^T + \delta\mathbf{I})^{-1}, \tag{2.37}$$

where $\mathbf{F}^{+}$ is the Moore-Penrose pseudoinverse of the matrix $\mathbf{F}$. This means, on one hand, that the posterior mean $\boldsymbol{\mu}_{\mathrm{post}}$ would be entirely determined by the data in the limit. On the other hand, in the limit, numerical evaluation of $\boldsymbol{\mu}_{\text{post}}$ would be difficult, since computation of the pseudoinverse is problematic when the matrix is ill-conditioned.

Improper priors

A prior that is not normalizable is called improper. Improper priors are widely used in Bayesian methodology, since the posterior can be normalizable even though the prior is not.

For example, a subset constraint of the form

$$p(\mathbf{x}) \propto \begin{cases} 1, & \text{if } \mathbf{x} \in S, \\ 0, & \text{otherwise,} \end{cases} \tag{2.38}$$

constitutes an improper prior, if $\int_S d\mathbf{x} = \infty$. This piecewise flat prior restricts the possible values of $\mathbf{x}$ to some subset S of the domain of definition of $p(\mathbf{x})$. When using this kind of a subset constraint as a prior, one should pay attention on what S is. If this subset is too large, the corresponding prior can be weak. As an example, a completely flat prior $p(\mathbf{x}) \equiv 1$ is weak. A subset constraint as a prior can be useful, for example, when the inverse problem needs to be solved only in a relatively small region of interest (ROI).

2.3.4 Posterior exploration

In addition to the construction of one or more posterior densities, another goal in the use of Bayesian methodology is to make inferences on $\mathbf{x}$ based on posterior probabilities. There exist a variety of different ways to evaluate and compare posteriors as well as to compute and evaluate estimates based on a single posterior.

Estimation of an unknown parameter requires exploration of a posterior. This work considers *maximum a posteriori* (MAP) and *conditional mean* (CM) estimates, which refer to numerical estimates of

$$\mathbf{x}_{\text{MAP}} = \arg\max_{\mathbf{x}} p(\mathbf{x} \mid \mathbf{y}) \quad \text{and} \quad \mathbf{x}_{\text{CM}} = \int \mathbf{x}\, p(\mathbf{x} \mid \mathbf{y})\, d\mathbf{x}, \tag{2.39}$$

respectively. Assuming that these are well defined finding either one can be a computationally challenging problem that requires the use of advanced optimization and numerical integration algorithms.

Difficulties arise whenever the shape of the posterior distribution is such that the algorithms tend to proceed to peculiar directions or get stuck into local minima. These difficulties can be caused e.g. by the general ill-posed and ill-conditioned nature of inverse problems, multimodality of the posterior, weakness of the prior density, and complexity of the forward model.

Once an estimate has been computed, its quality needs to be analyzed. Again, there are many applicable approaches. One strategy is simply to qualitatively compare the estimate to other estimates yielded by other models, e.g. using other prior densities. Another strategy is to compute credible (or credibility) intervals, i.e. intervals in which the unknown parameter lies with some given posterior probability values. A credible interval can be centered, for instance, at the posterior mean, or at the point corresponding to the narrowest possible interval, that is, the highest density interval. Third possible strategy is to visualize and analyze the convergence or behavior of the applied posterior exploration algorithm. Other strategies exist as well, including sophisticated methods of model comparison such as the use of Bayes factors [38], but these are not studied in this work.

MAP estimation

A comparison of the equations (2.3) and (2.23) reveals that finding a MAP estimate is closely connected with regularization; the same optimization methods that can be used to find the minimizer of (2.3) are applicable also for maximizing the argument of the exponential function in (2.23). Therefore, methods that are frequently used in MAP estimation include, for example, quasi-Newton and CG iterations.

For some distributions, like Gaussian distributions, the vectors $\mathbf{x}_{\mathrm{MAP}}$ and $\mathbf{x}_{\mathrm{CM}}$ coincide, but in general, they are distinct. It is often computationally less demanding to produce an estimate of $\mathbf{x}_{\mathrm{MAP}}$ than of $\mathbf{x}_{\mathrm{CM}}$, since the latter procedure requires multidimensional numerical integration. However, it is very typical that $\mathbf{x}_{\mathrm{MAP}}$ is less robust with respect to perturbation of the posterior. Namely, even a relatively small perturbations of the posterior can considerably change the maximizer of a probability density whereas the mean, which can be interpreted as the center of the probability mass, is more likely to stay relatively stable under perturbations.

Since MAP estimation algorithms are concentrated solely on finding the maximizer, they typically do not generate a representative sample from the posterior distribution. Therefore, based on a MAP estimation process, it can be difficult to infer, what the credibility of the estimate obtained is, or what the global characteristics of the posterior are. For these reasons, if the posterior is explored only through MAP estimation, it can remain unclear, e.g. whether the posterior is uni- or multimodal, or whether the prior is weak.

CM estimation

The central challenge in conditional mean estimation is that it necessitates multi-dimensional numerical integration. Traditional quadratures, such as Gaussian quadratures and Simpson's rule, do not scale well as the number n of the components of $\mathbf{x}$ increases. For example, according to the error analysis of the traditional quadratures, in numerical integration over the n-dimensional unit square $[0,1]^n$ using a regular grid, one would have to evaluate the integrand $\propto N^n$ grid points in order to achieve an estimate of accuracy $\propto 1/N^\rho$ with ρ being a quadrature specific constant. For this reason, MCMC sampling based numerical integration methods, in which the rate of convergence, in principle, does not depend on n, are extensively used in CM estimation.

MCMC sampling methods are algorithms that enable exploration of an arbitrary probability distribution. An MCMC algorithm produces a Monte Carlo estimate [33] of the conditional mean,

$$\frac{1}{N}(\mathbf{x}^{(1)} + \cdots + \mathbf{x}^{(N)}) \approx \int \mathbf{x}\, p(\mathbf{x} \mid \mathbf{y})\, d\mathbf{x}, \tag{2.40}$$

where $\{\mathbf{x}^{(i)}\}_{i=1}^{\infty}$ is an ergodic Markov chain with the posterior $p(\mathbf{x} \,|\, \mathbf{y})$ as the invariant density. Convergence of the estimate follows from the law of large numbers and the central limit theorem for ergodic Markov chains [37]. According to the central limit theorem, the estimation error is asymptotically Gaussian with the covariance matrix proportional to $1/N$.

The goal in MCMC sampling is to produce an ergodic Markov chain such that a reasonable convergence rate is achieved in practice. The rate of convergence in terms of central processing unit (CPU) time is affected by several factors, such as the complexity of the applied sampling strategy.

Even though consecutive sample points produced by the Markov Chain can be heavily correlated, points far from each other can be regarded as mutually independent and distributed according to the posterior. Once a sample has been generated, it not only enables CM estimation through (2.40), but it can be applied for error analysis of the CM estimate as well. Namely, based on the sample, one can estimate e.g. credibility intervals, marginal densities of the components of $\mathbf{x}$, as well as the sample quality.

There are several MCMC algorithms, such as the Metropolis-Hastings algorithm and the Gibbs sampler, that in general work well with slight adaptation. Given a posterior $p(\mathbf{x} \mid \mathbf{y})$ together with a proposal probability density $f(\mathbf{x}, \mathbf{z})$ satisfying $f(\mathbf{x}, \mathbf{z}) > 0$ if and only if $f(\mathbf{z}, \mathbf{x}) > 0$, the Metropolis-Hastings algorithm can be written into the following form:

- Given the current state $\mathbf{x}^{(k)}$, draw $\mathbf{z}$ from the proposal distribution determined by $\mathbf{z} \to f(\mathbf{x}^{(k)}, \mathbf{z})$.
- Draw $u \sim \text{Uniform}[0, 1]$ and update

  $$\mathbf{x}^{(k+1)} = \begin{cases} \mathbf{z}, & \text{if } u \leq r(\mathbf{x}^{(k)}, \mathbf{z}), \\ \mathbf{x}^{(k)}, & \text{otherwise}, \end{cases} \tag{2.41}$$

  where

  $$r(\mathbf{x}, \mathbf{z}) = \min\left\{1, \frac{p(\mathbf{z} \mid \mathbf{y}) f(\mathbf{z}, \mathbf{x})}{p(\mathbf{x} \mid \mathbf{y}) f(\mathbf{x}, \mathbf{z})}\right\}. \tag{2.42}$$

- Proceed to the next state ($k \to k + 1$).

Again, the simplest version of the Gibbs sampler, the systematic scan Gibbs sampler, is given by the following:

- Given the current state $\mathbf{x}^{(k)} = (x_1^{(k)}, \ldots, x_n^{(k)})$, for $i = 1, 2, \ldots, n$, draw $x_i^{(k+1)}$ from the conditional distribution

$$p(x_i \mid \mathbf{y}, x_1^{(k+1)}, \ldots, x_{i-1}^{(k+1)}, x_{i+1}^{(k)}, \ldots, x_n^{(k)}). \tag{2.43}$$

- Proceed to the next state ($k \to k+1$).

However, despite the simplicity of the algorithms, it is commonly agreed that producing a rapidly convergent sampling sequence is an art: Various factors, such as multimodality of the posterior and weakness of the applied prior, can slow down the convergence. Also, the computational complexity of the forward simulation is an important factor that affects the convergence rate in terms of CPU time. Namely, generation of a single sample point requires usually one or more evaluations of the forward mapping, and if the applied forward simulation is computationally heavy, sample generation can turn out to be a rather time-consuming process.

2.4 Electromagnetic biomedical applications

This section briefly reviews the electromagnetic applications studied in this work. Below, the mathematical notation slightly differs from what has been used above and there are some notational differences between the applications as well.

2.4.1 Electro/Magnetoencephalography

The EEG/MEG inverse problem [24] is to recover a primary current density $\mathbf{J}^p$ in an open two- or three-dimensional domain Ω, given an array of noisy external electric potential/magnetic field ($u/\mathbf{B}$) measurement data. The present inverse problem is linear; when discretized, the problem reduces to finding a vector $\mathbf{x}$ satisfying an underdetermined and ill-conditioned linear system of the form

$$\mathbf{y} = \mathbf{L}\mathbf{x} + \mathbf{n}, \tag{2.44}$$

where $\mathbf{L}$ is a so-called lead-field matrix, the vector $\mathbf{y}$ contains the measured data, and $\mathbf{n}$ is an error term caused by the noise in the measurements.

Forward model

In the standard EEG/MEG forward model, the electric field is assumed to be conservative, that is, $\mathbf{E} = -\nabla u$ and the total current density is given by

$$\mathbf{J} = \mathbf{J}^p + \mathbf{J}^s = \mathbf{J}^p - \sigma \nabla u, \tag{2.45}$$

where σ is the conductivity distribution of the target, and $\mathbf{J}^s = -\sigma \nabla u$ is the secondary or a volume current density. In this work, the CEM [49, 40] is applied for computation of u. The CEM assumes that contact electrodes $e_1, e_2, \ldots, e_L$ with contact impedances $z_1, z_2, \ldots, z_L$ are attached on the boundary $\partial\Omega$. Electrodes are assumed to be present during both the electric and magnetic field measurements. The electrode potentials form a voltage vector $\mathbf{U} = (U_1, U_2, \ldots, U_L)$ and the electric potential field u satisfies the equation

$$\nabla \cdot (\sigma \nabla u) = \nabla \cdot \mathbf{J}^p, \quad \text{in } \Omega, \tag{2.46}$$

as well as the boundary conditions

$$\sigma \frac{\partial u}{\partial n}\Big|_{\partial\Omega \setminus \cup_\ell e_\ell} = 0, \quad \int_{e_\ell} \sigma \frac{\partial u}{\partial n}\, ds = 0, \quad \text{and} \quad \Big(u + z_\ell \sigma \frac{\partial u}{\partial n}\Big)\Big|_{e_\ell} = U_\ell, \tag{2.47}$$

with $\ell = 1, 2, \ldots, L$. Additionally, the Kirchoff's voltage law $\sum_{\ell=1}^{L} U_\ell = 0$ is assumed to hold. The weak form [49, 18] of (2.46) and (2.47) can be formulated by requiring that $u \in H^1(\Omega) = \{\, w \in L^2(\Omega) \, : \, \partial w / \partial r_i \in L^2(\Omega),\, i = 1, 2, 3 \,\}$ and $\mathbf{J}^p \in \mathbf{H}(\text{div}; \Omega) = \{\, \mathbf{w} \in L^2(\Omega)^3 \, : \, \nabla \cdot \mathbf{w} \in L^2(\Omega) \,\}$. The space $H^1(\Omega)$ is the standard Sobolev space [9] and $\mathbf{H}(\text{div}; \Omega)$ [1, 35] is a space consisting of functions with a square integrable divergence. Note that the dipole current density [24] does not have a square integrable divergence.

The magnetic field point values, that are measured in MEG, can be obtained through a straightforward differentiation and integration procedure given the electric potential. Namely, the Biot-Savart law [26] combined with (2.45) states that the magnetic field at the point $\mathbf{r}_0$ outside the target Ω is given by the formula

$$\mathbf{B}(\mathbf{r}_0) = \frac{\mu_0}{4\pi} \int_\Omega (\mathbf{J}^\mathbf{p} - \sigma \nabla u) \times \frac{\mathbf{r}_0 - \mathbf{r}}{|\mathbf{r}_0 - \mathbf{r}|^3}\, d\mathbf{r}. \tag{2.48}$$

Forward simulation

In this work, the discretized fields $u_\mathcal{T}$ and $\mathbf{J}^p_\mathcal{T}$ corresponding to $u \in H^1(\Omega)$ and $\mathbf{J}^p \in \mathbf{H}(\text{div}; \Omega)$ and their coordinate vectors are defined by

$$u_\mathcal{T} = \sum_{i=1}^{N_u} \zeta_i \psi_i, \quad \text{and} \quad \mathbf{J}^p_\mathcal{T} = \sum_{i=1}^{N_J} x_i \, \boldsymbol{\psi}_i, \tag{2.49}$$

as well as $\boldsymbol{\zeta} = (\zeta_1, \zeta_2, \ldots, \zeta_{N_u})^T$ and $\mathbf{x} = (x_1, x_2, \ldots, x_{N_J})^T$, respectively. Here, the functions $\psi_1, \psi_2, \ldots, \psi_{N_u} \in H^1(\Omega)$ and $\boldsymbol{\psi}_1, \boldsymbol{\psi}_2, \ldots, \boldsymbol{\psi}_{N_J} \in \mathbf{H}(\text{div}; \Omega)$ are scalar and vector valued finite element basis functions, respectively, defined on a finite element mesh $\mathcal{T}$ [9]. The number of basis functions (number of degrees of freedom) is denoted by N_u and N_J, respectively. How the basis functions can be constructed in practice, is described e.g. in [1].

Furthermore, since in the CEM the sum of the electrode potentials is assumed to be zero, it is postulated that the finite element approximation of the electrode potential vector is given by

$$\mathbf{U}_\mathcal{T} = \mathbf{R}\boldsymbol{\xi}, \tag{2.50}$$

where $\boldsymbol{\xi}$ is an auxiliary vector and $\mathbf{R} \in \mathbb{R}^{L\times(L-1)}$ is a matrix with entries given by $R_{1,j} = -R_{j+1,j} = 1$ for $j = 1, 2, \ldots, L-1$, and $R_{i,j} = 0$ otherwise. The vectors $\mathbf{x}, \boldsymbol{\zeta}$, and $\boldsymbol{\xi}$ are linked through the linear system

$$\begin{pmatrix} \mathbf{A} & \mathbf{C} \\ \mathbf{C}^T & \mathbf{G} \end{pmatrix} \begin{pmatrix} \boldsymbol{\zeta} \\ \boldsymbol{\xi} \end{pmatrix} = \begin{pmatrix} \mathbf{F}\mathbf{x} \\ \mathbf{0} \end{pmatrix}, \tag{2.51}$$

where the submatrix entries are given by

$$F_{i,k} = \int_\Omega (\nabla \cdot \boldsymbol{\psi}_k)\psi_i \, d\Omega, \tag{2.52}$$

$$A_{i,j} = \int_\Omega \sigma \nabla\psi_i \cdot \nabla\psi_j \, d\Omega + \sum_{\ell=1}^{L} \frac{1}{z_\ell} \int_{e_\ell} \psi_i \psi_j \, dS, \tag{2.53}$$

$$C_{i,j} = -\frac{1}{z_1}\int_{e_1} \psi_i \, dS + \frac{1}{z_{j+1}} \int_{e_{j+1}} \psi_i \, dS, \tag{2.54}$$

$$G_{i,j} = \frac{1}{z_j}\int_{e_j} dS + \frac{\delta_{i,j}}{z_{j+1}} \int_{e_{j+1}} dS, \tag{2.55}$$

with $\delta_{i,j}$ denoting the Kronecker delta [9]. The system (2.51) arises from the Ritz-Galerkin discretization [9] of the weak form of (2.46) and (2.47). Similarly, a discretized version of (2.48) can be expressed as

$$\mathbf{B} = \mathbf{W}\mathbf{x} - \mathbf{V}\boldsymbol{\zeta}, \tag{2.56}$$

where $\mathbf{B}$ is a vector containing the magnetic field values at the measurement locations and the matrices are defined by

$$W_{i,3(j-1)+k} = \int_\Omega \frac{\mathbf{e}_k \cdot \boldsymbol{\psi}_j \times (\mathbf{r}_i - \mathbf{r})}{|\mathbf{r}_i - \mathbf{r}|^3} \, d\mathbf{r}, \tag{2.57}$$

$$V_{i,3(j-1)+k} = \int_\Omega \frac{\mathbf{e}_k \cdot \sigma \nabla\psi_j \times (\mathbf{r}_i - \mathbf{r})}{|\mathbf{r}_i - \mathbf{r}|^3} \, d\mathbf{r}, \tag{2.58}$$

with $\mathbf{r}_j$ denoting the jth measurement location and $\mathbf{e}_k$ denoting the kth natural basis vector.

The dependences of $\mathbf{U}_\mathcal{T}$ and $\mathbf{B}$ on the vector $\mathbf{x}$ are described by the electric and magnetic lead field matrices $\mathbf{L}^\mathrm{e}$ and $\mathbf{L}^\mathrm{m}$, respectively. These are given by

$$\mathbf{L}^\mathrm{e} = \mathbf{R}(\mathbf{C}^T\mathbf{B}^{-1}\mathbf{C} - \mathbf{G})^{-1}\mathbf{C}^T\mathbf{B}^{-1}\mathbf{F}, \tag{2.59}$$

$$\mathbf{L}^\mathrm{m} = \mathbf{W} - \mathbf{V}(\mathbf{B} - \mathbf{C}\mathbf{G}^{-1}\mathbf{C}^T)^{-1}\mathbf{F}, \tag{2.60}$$

as can be verified through straightforward linear algebra [19]. The formula for $\mathbf{L}^\mathrm{e}$ can be found by solving (2.51) with respect to $\boldsymbol{\xi}$ and multiplying the result with $\mathbf{R}$ to obtain $\mathbf{U}_\mathcal{T}$ as defined in (2.50). Again, the matrix $\mathbf{L}^\mathrm{m}$ can be obtained by solving (2.51) with respect to $\boldsymbol{\zeta}$ and substituting the result to (2.56). Note that the expressions (2.59) and (2.60) are valid only if a set of contact electrodes is attached to the head during the magnetic field measurement.

2.4.2 Electrical impedance tomography

In the present model for EIT, a current pattern $\mathbf{I} = (I_1, I_2, \dots, I_L)$ is injected into the domain Ω through the contact electrodes $e_1, e_2, \dots, e_L$. The injected currents induce a potential field u in the domain and respective voltages $\mathbf{U} = (U_1, U_2, \dots, U_L)$ on the electrodes. The measurement data are gathered by injecting a set of linearly independent current patterns and measuring the corresponding electrode voltages. The conductivity distribution σ in Ω is to be reconstructed from these voltage measurements. This is a non-linear inverse problem.

Forward model

Considering again the CEM, the electrode voltage vector $\mathbf{U}$ induced by the current pattern $\mathbf{I}$ can be found by solving the elliptic boundary value problem described by the equation

$$\nabla \cdot (\sigma \nabla u) = 0 \tag{2.61}$$

in the domain Ω, by the boundary conditions

$$\sigma \frac{\partial u}{\partial n}\Big|_{\partial\Omega \setminus \cup e_\ell} = 0, \quad \int_{e_\ell} \sigma \frac{\partial u}{\partial n}\, ds = I_\ell, \quad \text{and} \quad \left(u + z_\ell \sigma \frac{\partial u}{\partial n}\right)\Big|_{e_\ell} = U_\ell, \tag{2.62}$$

on $\partial\Omega$, with $\ell = 1, 2, \dots, L$, as well as by Kirchoff's current and voltage laws $\sum_{\ell=1}^{L} I_\ell = 0$ and $\sum_{\ell=1}^{L} U_\ell = 0$. According to [49], with certain assumptions made on the domain and on the conductivity distribution, there exists a unique pair $u \in H^1(\Omega)$ and $\mathbf{U} \in \mathbb{R}^L$ that solve this problem in the weak sense.

Forward simulation

The finite element approximations of u and $\mathbf{U}$, found as solutions of the CEM, are given by

$$u_{\mathcal{T}} = \sum_{i=1}^{N_u} \zeta_i \psi_i, \quad \text{and} \quad \mathbf{U}_{\mathcal{T}} = \mathbf{R}\boldsymbol{\xi}, \tag{2.63}$$

where the coefficients are obtained by solving the linear system

$$\begin{pmatrix} \mathbf{B} & \mathbf{C} \\ \mathbf{C}^T & \mathbf{G} \end{pmatrix} \begin{pmatrix} \boldsymbol{\zeta} \\ \boldsymbol{\xi} \end{pmatrix} = \begin{pmatrix} \mathbf{0} \\ \mathbf{R}^T \mathbf{I} \end{pmatrix} \tag{2.64}$$

similar to (2.51). Note that in (2.64) the dependence of the unknown vector on σ is highly non-linear. This means that each time σ is updated, the linear system has to be solved again. Note also that the matrix in (2.64) is, as a symmetric and positive definite FEM system matrix, well-conditioned despite the fact that the inverse problem itself is ill-conditioned.

2.4.3 Limited-angle computerized tomography

Computerized tomography (CT) is a discretized version of X-ray tomography [15, 14]. This work considers the linear inverse problem of the two-dimensional limited angle X-ray tomography, where the goal is to reconstruct an absorption distribution $f : \mathbb{R}^2 \to \mathbb{R}$ from its tomographic projections (radiographs), which constitute the measurement data.

Forward model

The forward model of two-dimensional limited-angle X-ray tomography is described by the Radon transform

$$\mathcal{R}f(t,\theta) = \int_{\mathbb{R}^2} f(x_1,x_2)\delta(x_1\cos\theta + x_2\sin\theta - t)\,dx_1dx_2, \tag{2.65}$$

where $t \in \mathbb{R}$, δ is Dirac's delta function [18] and the projection angle θ is limited to some sub-interval of $[-\pi/2, \pi/2)$. The Radon transform is linear and the value of $\mathcal{R}f(t,\theta)$ represents an integral of the absorption distribution over a straight line parametrized by t and the projection angle θ.

Forward simulation

In two-dimensional limited-angle CT, the data and solution spaces are pixelized, the number of radiographs is finite, each radiograph is approximated as a weighted sum of pixel values, and the data are assumed to contain additive noise. The resulting measurement model can be written as a linear system of the form

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{n}, \tag{2.66}$$

where $\mathbf{x}$ is the unknown consisting of the pixel values of the discretized f as its elements, $\mathbf{y}$ is the measurement data containing noisy readings of the discrete approximations of the integrals (2.65), $\mathbf{A}$ is the matrix corresponding to the discretized Radon transform, and $\mathbf{n}$ is the error term representing the measurement noise.

2.5 Forward simulation through p-FEM

The development of effective EIT and EEG/MEG forward simulations through application of h-FEM and p-FEM for electromagnetic fields is one of the main concerns of this work. The difference between these two versions of the FEM is in how the number of degrees of freedom is extended; in an h-version extension, the polynomial order p of the basis functions is relatively low and the element size h is reduced, whereas in the p-version, the polynomial order is increased and the mesh size is kept constant. The motivation for the use of p-FEM in addition to the traditional h-FEM in EIT and EEG/MEG forward simulations is that the solution of the CEM equations can be smooth in the interior domain and that the p-version is typically very efficient for problems with smooth solutions. The shape function (element basis function) types applied in this work for p-FEM forward simulation are briefly reviewed in this section considering both $H^1(\Omega)$- and $\mathbf{H}(\mathrm{div};\Omega)$-conforming shape functions.

2.5.1 Constant shape functions

In EIT and EEG/MEG, the conductivity distribution is to be approximated via the FEM. The only requirement for σ is that it must be square integrable, i.e. an $L^2(\Omega)$-function. Therefore, one can choose σ to be constant over each element of the finite element mesh, which results into globally piecewise constant conductivity distribution. This choice is used for its simplicity in this work; the simplest possible shape function is the one that is constant.

2.5.2 Lowest order Raviart-Thomas shape functions

In EEG/MEG, the approximation of the primary current density $\mathbf{J}^p_{\mathcal{T}} \in \mathbf{H}(\mathrm{div};\Omega)$ is represented using the lowest order Raviart-Thomas elements [9]; for each tetrahedral element, the ith shape function $\boldsymbol{\varphi}_i$ is linear, vanishes precisely at one of the ith vertex of the element, and the direction of $\boldsymbol{\varphi}_i$ is constant and given by the normal vector of the face opposite to the ith vertex (Figure 2.3). As a result, the number of degrees of freedom per element is four and the global finite element space is a subspace of $\mathbf{H}(\mathrm{div};\Omega)$, i.e. this element is $\mathbf{H}(\mathrm{div};\Omega)$-conforming.

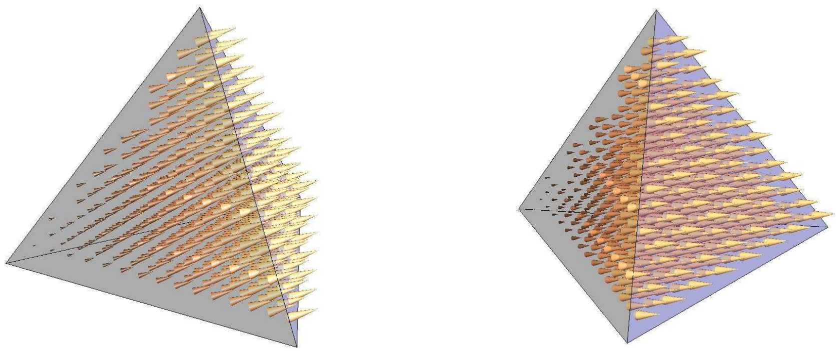

Figure 2.3: Two visualizations of a lowest-order Raviart-Thomas shape function in a tetrahedral element. The vector field in the tetrahedron is illustrated by the cones.

2.5.3 Hierarchic shape functions

The computation of the discretized electric potential field $u_{\mathcal{T}} \in H^1(\Omega)$ through p-FEM requires the use of higher-order elements, i.e. elements with shape functions of relatively high polynomial degree. Because numerical instability can occur in connection with higher-order elements, it is generally agreed that the shape functions should be hierarchic, i.e., the shape functions of maximal degree p are also shape functions of maximal degree $p+1$, and moreover, the number of shape functions not vanishing at the vertices and the sides of the elements is minimal. Hierarchic shape functions are constructed using Legendre polynomials

$$\mathsf{p}_k(x) = \frac{1}{2^k k!} \frac{d^k}{dx^k}(x^2-1)^k, \quad k \geq 0, \tag{2.67}$$

and the lowest order hierarchic shape functions are defined simply as the standard linear or bilinear nodal shape functions used in the h-FEM.

One dimensional case

In the one-dimensional case, the reference element is the interval $[-1,1]$. For this element, the one-dimensional hierarchic shape functions of polynomial order p are defined as

$$\varphi_1(\xi) = \frac{1-\xi}{2}, \quad \varphi_2(\xi) = \frac{1+\xi}{2}, \quad \varphi_n(\xi) = \phi_{n-1}(\xi), \quad n{=}3{,}4{,}\ldots{,}p{+}1, \tag{2.68}$$

where ϕ_n is defined as $\phi_n(\xi) = \sqrt{n-1/2} \int_{-1}^{\xi} \mathsf{p}_{n-1}(\xi)\, dt$. The shape functions are organized into two categories. The first one is formed by the polynomials φ_1 and φ_2, that are called the nodal shape functions, the external shape functions, or the vertex modes. The higher order polynomials $\varphi_3, \varphi_4, \ldots, \varphi_{p+1}$ form the second category. These vanish at the endpoints of the interval $[-1,1]$ and they are called the bubble functions, the internal shape functions, or the internal modes.

Quadrilateral elements

The two-dimensional quadrilateral reference element is the square $[-1,1] \times [-1,1]$. The corresponding two-dimensional hierarchical shape functions of polynomial order p are formed as products of one-dimensional shape functions as follows:

$$\varphi_{nm}(\xi,\eta) = \frac{1}{4}(1+(-1)^n\xi)(1+(-1)^m\eta), \quad n{=}1{,}2, \quad m{=}1{,}2, \tag{2.69}$$

$$\varphi_{nm}^{(0)}(\xi,\eta) = \phi_n(\xi)\phi_m(\eta), \quad n{=}2{,}3{,}\ldots{,}p, \quad m{=}2{,}3{,}\ldots{,}p, \tag{2.70}$$

$$\varphi_n^{(1)}(\xi,\eta) = \frac{1}{2}(1-\eta)\phi_n(\xi), \quad n{=}2{,}3{,}\ldots{,}p, \tag{2.71}$$

$$\varphi_n^{(2)}(\xi,\eta) = \frac{1}{2}(1-\xi)\phi_n(\eta), \quad n{=}2{,}3{,}\ldots{,}p. \tag{2.72}$$

These are organized to three categories: vertex modes φ_{nm}, internal modes $\varphi_{nm}^{(0)}$, and side modes $\varphi_n^{(1)}$, $\varphi_n^{(2)}$. In this work, only quadrilateral elements are used in the two-dimensional case.

Tetrahedral elements

In order to make indexing of three dimensional hierarchic shape functions convenient, this work adopts the notation [1] in which a tetrahedron with oriented edges is denoted by $\mathbf{t} = [t_1, t_2, t_3, t_4]$ with t_1, t_2, t_3, and t_4 indexing the vertices. Again, the sets containing the vertex, edge, and face indices of $\mathbf{t}$ are denoted by $\mathcal{V}(\mathbf{t})$, $\mathcal{E}(\mathbf{t})$, and $\mathcal{F}(\mathbf{t})$, respectively, and can be written explicitly as: $\mathcal{V}(\mathbf{t}) = \{t_1, t_2, t_3, t_4\}$, $\mathcal{E}(\mathbf{t}) = \{[t_1, t_2], [t_1, t_3], [t_1, t_4], [t_2, t_3], [t_2, t_4], [t_3, t_4]\}$, and $\mathcal{F}(\mathbf{t})$=$\{[t_1, t_2, t_3], [t_1, t_2, t_4], [t_1, t_3, t_4], [t_2, t_3, t_4]\}$.

The hierarchic shape functions for tetrahedral elements can be constructed using two reference tetrahedra $\hat{\mathbf{t}}_1 = [1, 2, 3, 4]$ (Figures 2.4) and $\hat{\mathbf{t}}_2 = [1, 3, 2, 4]$ (Figure 2.5) as proposed in [1]. Each of these share the same vertices $\hat{\mathbf{v}}_1 = (-1, 0, 0)$, $\hat{\mathbf{v}}_2 = (1, 0, 0)$, $\hat{\mathbf{v}}_3 = (0, \sqrt{3}, 0)$, and $\hat{\mathbf{v}}_4 = (0, 1/\sqrt{3}, 2\sqrt{2}/\sqrt{3})$, but their edges and faces are oriented differently. The reason for the use of more than one standard reference element is to minimize the difficulty of enforcing the conformity (global continuity), which can be problematic with hierarchic shape functions. Any tetrahedron with oriented edges in an unstructured tetrahedral finite element mesh can be reduced to either of the present two reference tetrahedra through cyclic rotations of the first three or the last three vertex indices, which are equivalent to rotations of two different faces [1].

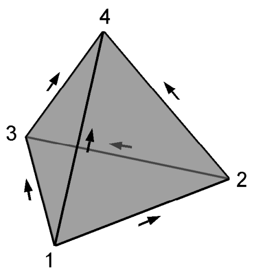

Figure 2.4: Reference tetrahedron $\hat{\mathbf{t}}_1 = [1, 2, 3, 4]$.

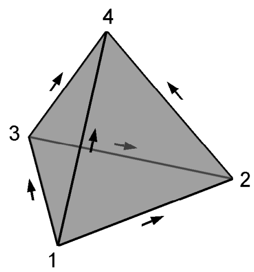

Figure 2.5: Reference tetrahedron $\hat{\mathbf{t}}_2 = [1, 3, 2, 4]$.

Hierarchic shape functions of $H^1(\Omega)$-conforming tetrahedral elements are divided into vertex, edge, face, and internal functions. The first three function types vanish at all but one of the vertices, edges, or faces, respectively, and the internal functions vanish on element boundaries. The vertex functions are first degree polynomials, and the other function types, also known as bubbles, are of higher degree.

The vertex functions of a tetrahedron $\mathbf{t} = [t_1, t_2, t_3, t_4]$ are defined as $\varphi^{(v)} = \lambda_v$ for $v \in \mathcal{V}(\mathbf{t})$, where λ_i denotes a barycentric coordinate, i.e. it is a linear polynomial that assumes the value one at the vertex i and vanishes on the opposite face. The corresponding edge, face, and internal modes of maximal degree p are defined as

$$\begin{aligned}
\varphi_j^{(\mathbf{e})} &= \lambda_{e_1}\lambda_{e_2}\mathsf{p}_j(\lambda_{e_2}-\lambda_{e_1}), && (2.73)\\
\varphi_{k_1k_2}^{(\mathbf{f})} &= \lambda_{f_1}\lambda_{f_2}\lambda_{f_3}\mathsf{p}_{k_1}(\lambda_{f_2}-\lambda_{f_1})\mathsf{p}_{k_2}(\lambda_{f_3}-\lambda_{f_1}), && (2.74)\\
\varphi_{\ell_1\ell_2\ell_3}^{(\mathbf{t})} &= \lambda_{t_1}\prod_{d=1}^{3}[\lambda_{t_{d+1}}\mathsf{p}_{\ell_d}(\lambda_{t_{d+1}}-\lambda_{t_1})] && (2.75)
\end{aligned}$$

respectively. Here, the superscript $\mathbf{e}$ denotes an edge $[e_1, e_2] \in \mathcal{E}(\mathbf{t})$ and $\mathbf{f}$ denotes a face $[f_1, f_2, f_3] \in \mathcal{F}(\mathbf{t})$, and the subscripts determining the Legendre polynomial degrees are positive or zero integers such that $j \leq p-2$, $k_1+k_2 \leq p-3$, and $\ell_1+\ell_2+\ell_3 \leq p-4$.

A scalar shape function φ supported on $\mathbf{t}$ and the corresponding reference element shape function $\hat{\varphi}$ are related through the transformation $\varphi = \hat{\varphi}\circ\mathbf{F}_{\mathbf{t}}^{-1}$, where $\mathbf{F}_{\mathbf{t}}$ is an invertible affine mapping from the corresponding reference tetrahedron onto $\mathbf{t}$. Using this transformation, the $H^1(\Omega)$-conforming hierarchic basis functions for unstructured tetrahedral meshes can be constructed in the usual way using the reference element shape functions.

2.5.4 Practical aspects of implementation

There are a few differences in implementation of h- and p-version FEM solvers on a general level. A central difference is that the construction of the system matrix is computationally more requiring in p-FEM; due to the relatively high polynomial order of the basis functions, high-order Gaussian quadratures [50] are needed when computing the integrals of the p-FEM system matrix. Moreover, the percent of the resulting system matrix non-zero entries is higher due to the large number of p-FEM basis functions belonging to one element.

Since the p-FEM lies on the use of Legendre polynomials, point evaluation of the basis functions is not as straightforward as in the case of h-FEM. An efficient algorithm to evaluate Legendre polynomials and their derivatives at a given point can be constructed, e.g. through the recursion

$$(n+1)\mathsf{p}_{n+1}(x) = (2n+1)x\mathsf{p}_n(x) - n\mathsf{p}_{n-1}(x) \tag{2.76}$$

based iterations [10] or by computing all the values at the quadrature points in advance.

3 Results

This chapter summarizes the results and findings of the papers [**I**]–[**VI**].

3.1 [I]

This study investigated EIT and the two-stage reconstruction process, in which a ROI is determined in the first stage and the actual reconstruction is found in the second stage through MCMC sampling. This two-stage approach was found to be applicable for detection of circular anomalies. An effective FEM based smoothness prior was discovered. An anomaly prior constructed using a four-dimensional parametrization of the EIT inverse problem together with a piecewise constant interpolation scheme was successfully applied. The performance of the MCMC was found to be superior when the FEM was used in the forward simulation as opposed to the linearized approximation. Additionally, an enhanced noise model incorporating *a priori* information about the forward simulation error into the measurement error model was found to be superior over the white noise model when the linearized FEM forward simulation was used. This study also revealed a need to study more sophisticated MCMC sampling as well as more rigorous forward simulation techniques.

3.2 [II]

In this study, the p-FEM was applied to EIT forward simulation as motivated by [**I**]. It was shown numerically by using the unit square as the computational domain that the performance of the p-FEM is better than that of the h-version FEM when uniform mesh refinement is used. It was also found that the characteristic difference between the h- and p-versions of the FEM is that a lot more computational effort is spent on construction of a system matrix in the p-version. Consequently, finding computationally tractable ways to obtain system matrices needed in EIT reconstruction, e.g. through the use of *a priori* information, is mentioned as a future work topic. Based on the results, as a potential target for future work, the paper also proposes the hp-FEM, in which both the mesh size and the polynomial degree of the shape functions can be varied simultaneously.

3.3 [III]

This study concerned again EIT and conductivity anomalies. An efficient reconstruction algorithm based on the Kirsch factorization method was presented for two-dimensional polygonal domains. Through numerical experiments, in which the p-FEM forward simulation developed for [**II**] was used, the algorithm was shown to provide information about shapes and locations for a wide class of conductivity anomalies. Both circular and kite-shaped anomalies were successfully recovered in the numerical experiments.

As a future work topic, the development of a systematic method for taking the measurement noise into account is proposed. It is also mentioned that a study of potentially rapid computational forward simulation methods would be needed for the future development of the algorithm.

3.4 [IV]

This study compared the performances of h- and p-type finite elements in EEG/MEG forward simulation. Both forward simulation types were described and implemented based on the complete electrode model. The performance comparison was done by analyzing direct measurement of the relative discretization error as well as via reconstruction of multiple dipole-like electric sources applying the regularized FOCUSS algorithm. It was found that in terms of the relative discretization error the performance of the p-type FEM was superior to that of the h-type. In contrast, the comparison of the reconstructions revealed only small differences.

Mesh generation for the p-FEM, based on the *a priori* knowledge of what is a good choice for a dipole-like source diameter, is proposed as a topic for future work. Another topic proposed for the future is to investigate whether p-FEM forward simulation provides a suitable way to recover both dipole moment and depth at the same time. Additionally, it is proposed that there is a need to study the hp-FEM and different element types, such as prism elements.

3.5 [V]

This study considered the MEG/EEG inverse problem through hierarchical Bayesian models, in which the variance of the primary current density was assumed to be distributed according to the generalized gamma hyperprior density. The paper proposes an iterative alternating sequential (IAS) algorithm for computing the MAP estimate simultaneously for the primary current density and its variance. The IAS algorithm was found to be very fast and easy to implement. It is also shown in the paper, that for particular choices of the scalar hyperparameters specifying the hyperprior, the algorithm effectively approximates popular regularization strategies. The MAP estimates produced by IAS were compared against MCMC based CM estimates. Like in the study of [**I**], a ROI was used in the computation of the CM estimates.

Both a planar geometry and a realistic geometry were used in the numerical experiments employing multiple dipole-like currents as sources. Forward simulation techniques developed in [**IV**] were applied in the case of the realistic geometry. The results suggest that the CM estimate is most effective in combination with the inverse gamma prior, and that the MAP estimate, on the other hand, is most effective when applied in connection with the gamma prior. The results also suggest that, in general, CM estimation is superior over MAP estimation of the focal activity in different parts of the brain.

As topics for future work, the paper proposes a hierarchical extension of the model, where the values of the hypermodel parameters would be chosen based on the data, as well as an extension of the formalism to include time dependent sources with a longitudinal correlation structure.

3.6 [VI]

This study considered linear inverse problems. It proposed and tested an iterative coarse-to-fine reconstruction procedure based on iterative regularization via the CG method, on Tikhonov preconditioning, as well as on wavelet low-pass filtering. The numerical experiments concerning this procedure showed that a combined use of wavelet filters and preconditioning can be effective in limited-angle CT reconstruction. The reconstructions obtained by using the proposed coarse-to-fine reconstruction procedure were superior to the reconstructions produced by other reconstruction strategies only when both wavelet filtering and preconditioning were applied simultaneously.

The paper suggests a future work to design sophisticated multiresolution filters that decompose the solution space to detectable and undetectable parts. Another future study suggested is an effort to develop a regularization technique, which would render the proposed coarse-to-fine strategy less sensitive to the measurement noise. Moreover, finding a reasonable method for choosing the adjustable parameters is mentioned as an important future work.

4 Discussion and conclusions

This work discusses computational methods in electromagnetic biomedical inverse problems. The applications include EIT, EEG/MEG and limited-angle CT. These correspond to inverse problems of sounding, source determination, and high-frequency imaging, respectively. Both classical regularization techniques and Bayesian methodology have been applied in solution of these inverse problems.

The focus of this work is more on development of a variety of case-by-case computational methods and techniques than of a single computational strategy that would cover all the applications. Papers [**II**]–[**V**] can be seen to be partially motivated by [**I**], which is chronologically the first one. Namely, the importance of effective forward simulation is emphasized in [**I**], and in [**II**]–[**V**], effective forward simulation strategies are developed and applied. However, also inversion methods are studied in these. Paper [**VI**] can be regarded as complementary, since it discusses linear inverse problems from a general point of view, a combined use of preconditioning and projections in iterative regularization, *a priori* information in general, as well as a forward model not requiring numerical simulation of Maxwell's equations.

The results obtained in this work suggest that the use of p-FEM in EIT and EEG/MEG forward simulations is advantageous as compared to the h-FEM. Furthermore, the overall experience from this work suggests that FEM, in general, can be regarded as favorable. The advantage of the FEM is that the system matrices, even though often large, are relatively sparse and the number of the systems to be solved depends on the amount of the gathered data, which for instance in EIT and EEG/MEG corresponds to the number of data channels.

Another important finding of this work is, that in EIT and EEG/MEG, two-stage reconstruction of anomalies utilizing Bayesian methodology can be advantageous, when the first-stage reconstruction is produced through MAP estimation and the second stage is computed employing MCMC integration together with a ROI. The resulting reconstruction process is not fully Bayesian, since it violates the likelihood principle according to which one should not use the data in the construction of the prior. However, both reconstruction stages are fully Bayesian when considered individually and, moreover, the likelihood principle itself is commonly agreed to be somewhat controversial.

There are other findings concerning Bayesian methodology as well, covering among others, construction of different prior densities. For instance in [**I**], a smoothness prior is constructed using a FEM based stiffness matrix, and a locally uninformative anomaly prior based on a specific parametrization of the inverse problem is introduced. Again, in [**V**], generalized gamma densities are used as a hyperpriors and the IAS algorithm is proposed for MAP estimation. Based on the results it is suggested in [**V**], that the CM estimate is preferable in combination with the inverse gamma prior while the

MAP estimate is preferable when the standard gamma prior is applied. Furthermore, the results of [**V**] suggest that the performance of CM estimation can be, in general, superior over that of MAP estimation.

Numerical results obtained through different regularization strategies are presented in [**III**], [**IV**], and [**VI**]. In [**III**], it is shown that the Kirsch factorization method can be successfully used in EIT for recovery of both convex and non-convex anomalies. In [**IV**], a regularized FOCUSS algorithm is used in evaluation of the forward simulation effectiveness. Again, in [**VI**], it is shown that a combined use of wavelet filtering and preconditioning can be necessary in iterative regularization of linear inverse problems.

Since progress has been made regarding both regularization and Bayesian methodology, there seems to be no need to concentrate the future work in just one of these approaches. Which one is the preferable depends largely on the target application as well as on the nature of the available *a priori* knowledge. However, as it has been shown in this work and elsewhere in the literature [29], a variety of regularization methods can be given a Bayesian interpretation, and therefore, it seems obvious that the interest towards the Bayesian approach to inverse problems [12, 25, 29] will continue to grow in the future.

In the future studies on the EIT and EEG/MEG forward simulations, the next step would be to perform numerical experiments using data obtained from real measurements to provide practical experience of forward simulation accuracy. It would also be interesting to apply p-FEM forward simulation to other biomedical imaging modalities involving diffusion (Poisson-type) equations like (2.46) and (2.61). These applications could include electrocardiography/magnetocardiography (ECG/MCG) [46, 47], in which the potential distribution of the heart is to be determined, as well as diffuse optical tomography (DOT) [4, 45], in which optical absorption properties of a body are estimated from the transmission and scattering data of near infrared light.

Considering the future work in general, an important aspect is that the computational power of personal computers has increased tremendously during the course of this work. It is currently possible to solve partial differential equations numerically in realistic three dimensional geometries using a regular PC. Consequently, simulation of realistic models has become feasible in many senses while rough simplifications, i.e. spherical cows [23], have become decreasingly necessary. The future work should, therefore, be as much as possible concentrated on the use of realistic or authentic models, geometries, and measurement data.

Obviously, development of algorithms that would take all the advantages of the growing computational power is a challenging task. However, with regard to regular personal computers, it is clear in a number of cases how some additional computational power could be used. For example, when MCMC sampling is applied to the EIT inverse problem, in the generation of each sample point, there is a need to solve one or more linear systems, the sizes of which are determined by the accuracy of the applied forward simulation. At the moment, the computation must either be restricted into a ROI, as has been done in this work, or the quality of the forward simulation must be compromised in order to achieve a reasonable sampling rate in practice. Implementation of a fast sampler in a realistic geometry without compromises can easily require a computer many times faster than the fastest existent personal computers. Again, the amount of measurement data gathered in EIT can easily be so large that there can be a need to make compromises due to practical limitations in computational capabilities. For example, if 100 electrodes are used together with 99 linearly independent current patterns, the length of the resulting data vector will be $99 \times 100 = 9900$ and the corresponding likelihood covariance matrix will be a 9900-by-9900 full matrix, which can currently be regarded as very large to be handled rapidly in a reconstruction process.

Bibliography

[1] M Ainsworth and J Coyle. Hierarchic finite elements for unstructured tetrahedral meshes. *Int. J. Numer. Meth. Engng*, 58:2103–2130, 2003.

[2] H Ammari, E Beretta, and E Francini. Reconstruction of thin conductivity imperfection. *Applicable Analysis*, 83:63–76, 2004.

[3] K E Andersen, S P Brooks, and M B Hansen. A bayesian approach to crack detection in electrically conducting media. *Inverse Problems*, 17:121–136, 2001.

[4] S R Arridge. Topical review: Optical tomography in medical imaging. *Inverse Problems*, 15(2):R41–R93, 1999.

[5] T Auranen, A Nummenmaa, M S Hämalainen, I P Jääskelainen, J Lampinen, A Vehtari, and M Sams. Bayesian analysis of the neuromagnetic inverse problem with l(p)-norm priors. *Neuroimage*, 26(3):870–884, 2005.

[6] S Babaeizadeh, D H Brooks, D Isaacson, and J Newell. Electrode boundary conditions and experimental validation for BEM-based EIT forward and inverse solutions. *IEEE Transactions on Medical Imaging*, 25(9):1180 – 1188, 2006.

[7] R Barbuzza, M Vénere, and A Clausse. Tomographic reconstruction using heuristic monte carlo methods. *Journal of Heuristics*, 13(3):227–242, 2007.

[8] E Beretta, E Francini, and M Vogelius. Asymptotic formulas for steady state voltage potentials in the presence of thin inhomogeneities. a rigorous error analysis. *Journal des Mathématiques Pures et Appliqùees*, 82(10):1277–1301, 2003.

[9] D Braess. *Finite Elements*. Cambridge University Press, Cambridge, 2001.

[10] M Byckling. Implementation of the *hp*-element method in elmer. Master's thesis, Helsinki University of Technology, 2005.

[11] A P Calderón. On an inverse boundary value problem. In *Seminar on Numerical Analysis and its Applications to Continuum Physics*, pages 65–73, Rio de Janeiro, 1980. Soc. Brasil. Mat.

[12] D Calvetti and E Somersalo. *Intorduction to Bayesian Scientific Computing — Ten Lectures on Subjective Computing*. Springer-Verlag, Berlin, 2007.

[13] M Cheney, D Isaacson, and J C Newell. Electrical impedance tomography. *SIAM Review*, 41:85–101, 1999.

[14] M E Davison. The ill-conditioned nature of the limited-angle tomography problem. *SIAM J. Appl. Math.*, 43:428–448, 1983.

[15] S Deans. *The Radon transform and some of its applications*. John Wiley & Sons, New Jersey, 1983.

[16] N Von Ellenrieder, C H Muravchik, and A Nehorai. MEG forward problem formulation using equivalent surface current densities. *IEEE Transactions on Biomedical Engineering*, 52(7):1210–1217, 2005.

[17] J Enderle. *Introduction to Biomedical Engineering, Second Edition*. Academic Press, London, 2005.

[18] L C Evans. *Partial differential equations*. American Mathematical Society, Rhode Island, 1998.

[19] G Golub and C van Loan. *Matrix Computations*. The John Hopkins University Press, Baltimore, 1989.

[20] I F Gorodnitsky and B D Rao. Sparse signal reconstruction from limited data using FOCUSS: a re-weighted minimum norm algorithm. *IEEE Trans. Signal Proc.*, 45:600–616, 1997.

[21] M Hanke. Iterative regularization techniques in image restoration. In D Colton et al, editor, *Surveys on Solution Methods for Inverse Problems*, pages 35–52, Wien, 2000. Springer-Verlag.

[22] K M Hanson and G W Wecksung. Bayesian approach to limited-angle reconstruction in computed tomography. *Journal of the Optical Society of America*, 73:1501–1509, 1983.

[23] J Harte. *Consider a Spherical Cow: A Course in Environmental Problem Solving*. University Science Books, Herndon, VA, USA, 1988.

[24] M Hämäläinen, R Hari, R J Ilmoniemi, J Knuutila, and O V Lounasmaa. Magnetoencephalography — theory, instrumentation, and applications to invasive studies of the working human brain. *Reviews of Modern Physics*, 65:413–498, 1993.

[25] J Idier. *Bayesian Approach to Inverse Problems*. Wiley-ISTE, New Jersey, 2008.

[26] J D Jackson. *Classical Electrodynamics, 3rd Edition*. John Wiley & Sons, New Jersey, 1998.

[27] K Jerbi, C Mosher, S Baillet, and R M Leahy. On MEG forward modelling using multipolar expansions. *Physics in Medicine & Biology*, 47:523–555, 2002.

[28] P Jinsong and H Suyi. Noisy temperature field reconstruction by wavelet-expansion in OCT measurement. *Heat and Mass Transfer*, 38(6):507–512, 2002.

[29] J P Kaipio and E Somersalo. *Statistical and Computational Methods for Inverse Problems*. Springer, Berlin, 2004.

[30] A Kirsch. Characterization of the shape of a scattering obstacle using the spectral data of the far field operator. *Inverse Problems*, 14:1489–1512, 1998.

[31] S Komssi, J Huttunen, H J Aronen, and R J Ilmoniemi. EEG minimum-norm estimation compared with MEG dipole fitting in the localization of somatosensory sources at S1. *Clinical Neurophysiology*, 115(3):534–542, 2004.

[32] J Liu, B Guerrier, and C Benard. A sensitivity decomposition for the regularized solution of inverse heat conduction problems by wavelets. *Inverse Problems*, 11:1177–1187, 1995.

[33] J S Liu. *Monte Carlo strategies in scientific computing*. Springer Verlag, New York, 2001.

[34] S G Mallat. A theory for multiresolution signal decomposition: The wavelet representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 7:674–693, 1989.

[35] P Monk. *Finite Element Methods for Maxwell's Equations*. Clarendon Press, Oxford, UK, 2003.

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

[37] E Nummelin. *General Irreducible Markov Chains and Non-negative Operators*. Cambridge University Press, Cambridge, 1984.

[38] A O'Hagan and F Forster. *Kendall's Advanced Theory of Statistics, Volume 2B: Bayesian Inference*. Arnold, London, 2004.

[39] G Olafsson and E T Quinto. *The Radon Transform, Inverse Problems, and Tomography*. American Mathematical Society, Boston, MA, USA, 2006.

[40] J Ollikainen, M Vauhkonen, P A Karjalainen, and J P Kaipio. Effects of electrode properties on EEG measurements and a related inverse problem. *Medical Engineering & Physics*, 22:535–545, 2000.

[41] C Phillips, M D Rugg, and K J Friston. Systematic regularization of linear inverse solutions of the EEG source localization problem. *NeuroImage*, 17:287–301, 2002.

[42] M Piana and M Bertero. Projected landweber method and preconditioning. *Inverse Problems*, 13:441–463, 1997.

[43] B D Rao, K Engan, S F Cotter, J Palmer, and K Kreutz-Delgado. Subset selection in noise based on diversity measure minimization. *IEEE Transactions on Signal Processing*, 51(3):760–770, 2003.

[44] C Schwab. *p- and hp- Finite Element Methods*. Oxford University Press, New York, 1998.

[45] M Schweiger, S R Arridge, and D T Delpy. Application of the finite-element method for the forward and inverse models in optical tomography. *Journal of Mathematical Imaging and Vision*, 3(3):263–283, 1993.

[46] A V Shahidi, P Savard, and R Nadeau. Forward and inverse problems of electrocardiography: modeling andrecovery of epicardial potentials in humans. *IEEE Transactions on Biomedical Engineering*, 41(3):249–256, 1994.

[47] G Shou, L Xia, M Jiang, F Liu, and S Crozier. *Functional Imaging and Modeling of the Heart*. Springer, Berlin, 2007.

[48] M Soleimani, C E Powell, and N Polydorides. Improving the forward solver for the complete electrode model in EIT using algebraic multigrid. *IEEE transactions on medical imaging*, 24(5):577–583, 2005.

[49] E Somersalo, M Cheney, and D Isaacson. Existence and uniqueness for electrode models for electric current computed tomography. *SIAM J. Appl. Math.*, 52:1023–1040, 1992.

[50] B Szabo and I Babuska. *Finite Element Analysis*. John Wiley & Sons, New York, 1991.

[51] O Tanzer, S Järvenpää, J Nenonen, and E Somersalo. Representation of bioelectric current sources using whitney elements in the finite element method. *Physics in Medicine and Biology*, 50:3023–3039, 2005.

[52] A N Tikhonov, A V Goncharsky, V V Stepanov, and A G Yagola. *Numerical Methods for the Solution of Ill-posed Problems*. Kluwer, Dordrecht, 1995.

[53] R Van Uitert, D Weinstein, and C R Johnson. Volume currents in forward and inverse magnetoencephalographic simulations using realistic head models. *Annals of Biomedical Engineering*, 31:21–31, 2003.

[54] R Van Uitert, D Weinstein, C R Johnson, and L Zhukov. Finite element EEG and MEG simulations for realistic head models: Quadratic vs. linear approximations. *Journal of Biomedizinische Technik*, 46:32–34, 2001.

[55] G Xu, H Wu, S Yang, S Liu, Y Li, Q Yang, W Yan, and M Wang. 3-D electrical impedance tomography forward problem with finite element method. *IEEE transactions on magnetics*, 41(5):1832 – 1835, 2005.

Part II

Original papers

(continued from the back cover)

HELSINKI UNIVERSITY OF TECHNOLOGY INSTITUTE OF MATHEMATICS
RESEARCH REPORTS

The reports are available at *http://math.tkk.fi/reports/* .
The list of reports is continued inside the back cover.

ISBN 978-951-22-9679-8 (print)
ISBN 978-951-22-9680-4 (PDF)
ISSN 0784-3143 (print)
ISSN 1797-5867 (PDF)
TKK Mathematics, 2008