Department of Biomedical Engineering and Computational Science

Regularization methods for diffuse optical tomography

Petri Hiltunen

Aalto University

DOCTORAL DISSERTATIONS

Aalto University publication series
DOCTORAL DISSERTATIONS 98/2011

Regularization methods for diffuse optical tomography

Petri Hiltunen

Doctoral dissertation for the degree of Doctor of Science in Technology to be presented with due permission of the School of Science for public examination and debate in Auditorium E at the Aalto University School of Science (Espoo, Finland) on the 8th of November 2011 at 12 noon.

Aalto University
School of Science
Department of Biomedical Biomedical and Computational Science

Supervisor
Professor Jouko Lampinen

Instructor
Professor Jouko Lampinen

Preliminary examiners
PhD Ville Kolehmainen, University of Eastern Finland, Finland
PhD Oliver Dorn, University of Manchester, UK

Opponent
Professor Colin Fox, University of Otago, New Zealand

Aalto University publication series
DOCTORAL DISSERTATIONS 98/2011

ISBN 978-952-60-4321-0 (pdf)
ISBN 978-952-60-4320-3 (printed)
ISSN-L 1799-4934
ISSN 1799-4942 (pdf)
ISSN 1799-4934 (printed)

Aalto Print
Helsinki 2011

Finland

The dissertation can be read at http://lib.tkk.fi/Diss/

A! Aalto University

Aalto University, P.O. Box 11000, FI-00076 Aalto www.aalto.fi

Abstract

Author
Petri Hiltunen

Name of the doctoral dissertation
Regularization methods for diffuse optical tomography

Publisher School of Science

Unit Department of Biomedical Engineering and Computational Science

Series Aalto University publication series DOCTORAL DISSERTATIONS 98/2011

Field of research Computational Science

Manuscript submitted 1 April 2011 **Manuscript revised** 9 June 2011

Date of the defence 8 November 2011 **Language** English

☐ **Monograph** ☒ **Article dissertation (summary + original articles)**

Abstract

Near-infrared light can be used as a three dimensional imaging tool, called diffuse optical tomography (DOT), in the study of human physiology. Due to differences in the extinction coefficients of oxygenated and deoxygenated haemoglobin at different wavelengths, concentrations of the haemoglobins can be resolved from measurements at a few wavelengths. Therefore, DOT is a fascinating modality for biomedical applications, such as functional brain imaging, breast cancer screening, etc. Moreover light is a safe tool, because it is non-ionizing and at intensity levels used in DOT, it does not cause burns at skin or in organs.

There are a few different models to describe light propagation in tissue-like media. One of the simplest, called the diffusion approximation (DA), was used in this thesis. The optical properties, the absorption and the scattering coefficients, are the parameters which determine the light propagation in the DA model. When optical properties are known and one is interested in estimating the photon flux at the boundary, the problem is called a forward problem. Likewise, when the photon flux at the boundary is measured and the task is to find the optical properties, the problem is called an inverse problem. The inverse problem related to DOT is ill-posed, i.e., the solution might not be unique or the solution does not depend continuously on data.

Due to ill-posedness of the inverse problem, some regularization methods should be used. In this thesis, regularization methods for a stationary and nonstationary inverse problems was considered. By the nonstationary inverse problem, it is meant that the optical properties are not static during the measurement and the whole evolution of the optical properties is reconstructed in contrast to the stationary problem, where the optical properties are assumed to be static during the measurement.

The regularization for the inverse problem could be implemented as the Tikhonov regularization or using statistical inversion theory, also known as the Bayesian framework. In this thesis, two different regularization methods for the static reconstruction problem in DOT were studied. They both allow discontinuities in the optical properties that might occur at boundaries between organs. For the nonstationary reconstruction problem, an efficient regularization model is presented.

Keywords diffuse optical tomography, inverse problem, regularization

ISBN (printed) 978-952-60-4320-3 **ISBN (pdf)** 978-952-60-4321-0

ISSN-L 1799-4934 **ISSN (printed)** 1799-4934 **ISSN (pdf)** 1799-4942

Location of publisher Espoo **Location of printing** Helsinki **Year** 2011

Pages 143 **The dissertation can be read at** http://lib.tkk.fi/Diss/

A! Aalto-yliopisto

Aalto-yliopisto, PL 11000, 00076 Aalto www.aalto.fi

Tiivistelmä

Tekijä
Petri Hiltunen

Väitöskirjan nimi
Regularisointimenetelmät diffuusissa optisessa tomografiassa

Julkaisija Perustieteiden korkeakoulu

Yksikkö Lääketieteellisen tekniikan ja laskennalisen tieteen laitos

Sarja Aalto University publication series DOCTORAL DISSERTATIONS 98/2011

Tutkimusala Laskennallinen tiede

Käsikirjoituksen pvm 01.04.2011 **Korjatun käsikirjoituksen pvm** 09.06.2011

Väitöspäivä 08.11.2011 **Kieli** Englanti

☐ **Monografia** ☒ **Yhdistelmäväitöskirja (yhteenveto-osa + erillisartikkelit)**

Tiivistelmä

Lähi-infrapunavaloa voidaan käyttää kolmiulotteiseen kuvantamiseen ihmiskehon fysiologian tutkimuksessa. Menetelmää kutsutaan diffuusiksi optiseksi tomografiaksi (DOT). Hapellisen ja hapettoman hemoglobiinin erilaisista valon vaimennuskertoimista eri aallon pituuksilla johtuen hemoglobiinien konsentraatiot pystytään määrittämään käyttämällä useampaa aallonpituutta mittauksissa. Sen takia DOT on mielenkiintoinen menetelmä lääketieteellisiin sovelluksiin, kuten aivojen funktionaaliseen kuvantamiseen, rintasyövän seurantaan jne. Lisäksi valo on turvallinen säteilyn muoto, koska säteily ei ole ionisoivaa eikä käytetty intensiteettitaso aiheuta palovammoja ihoon tai muihin elimiin.

Valon kulun mallintamiseen on esitetty muutamia malleja. Yhtä yksinkertaisimmista kutsutaan diffuusioapproksimaatioksi (DA), jota käytetään tässä työssä. Optiset ominaisuudet, erityisesti absorptio- ja sirontakerroin, ovat mallin parametreja, jotka määräävät valon etenemisen DA:ssa. Kun optiset ominaisuudet ovat tunnetut ja ollaan kiinnostuneita arvioimaan fotonivuo kohteen pinnalla, tehtävää kutsutaan suoraksi ongelmaksi. Kun fotonivuo pinnalla on mitattu ja tehtävänä on selvittää optiset ominaisuudet, tehtävää kutsutaan käänteisongelmaksi. Käänteisongelma DOT:ssa on heikosti aseteltu, mikä tarkoittaa sitä, että ratkaisu ei ole välttämättä yksikäsitteinen tai ratkaisu ei riipu jatkuvasti mittausdatasta.

Heikosti asetetusta käänteisongelmasta johtuen, on käytettävä jotakin regularisointimenetelmää. Tässä työssä on tutkittu regularisointia sekä DOT:in stationaarisessa että epästationaarisessa käänteisongelmassa. Epästationaarisella käänteisongelmalla tässä työssä tarkoitetaan sitä, että optiset parametrit muuttuvat mittauksen aikana ja optisten ominaisuuksien koko muutoshistoria rekonstruoidaan. Sitä vastoin stationaarisessa käänteisongelmassa optiset parametrit oletetaan muuttumattomiksi koko mittauksen ajan.

Regularisointi käänteisongelmaan voidaan toteuttaa esimerkiksi Tikhonov-regularisaationa tai käyttämällä tilastollisia käänteisongelman ratkaisumenetelmiä, joita kutsutaan myös Bayesilaisiksi menetelmiksi. Tässä työssä tutkittiin kahden erilaisen regularisaatiomenetelmän soveltuvuutta stationaariseen DOT:aan. Kummatkin menetelmät sallivat epäjatkuvuuksia optisissa parametreissa, mikä voi tapahtua kudosten rajapinnoilla. Epästationaariseen käänteisongelmaan on työssä esitetty tehokas regularisointimenetelmä.

Avainsanat diffuusi optinen kuvantaminen, käänteisongelma, regularisointi

ISBN (painettu) 978-952-60-4320-3 **ISBN (pdf)** 978-952-60-4321-0

ISSN-L 1799-4934 **ISSN (painettu)** 1799-4934 **ISSN (pdf)** 1799-4942

Julkaisupaikka Espoo **Painopaikka** Helsinki **Vuosi** 2011

Sivumäärä 143 **Luettavissa verkossa osoitteessa** http://lib.tkk.fi/Diss/

Preface

It is 25 September, 00:27 am and I am finalizing this thesis. This is the first time, and I hope the last time, when I am working in the night. This thesis will end my 22 years long journey in the Finnish education system from Kirkonkylän ala-aste to Aalto University. It does not mean that I know everything but rather it is just the beginning for a lifetime of learning.

My journey to optical imaging began in 2003 in the Laboratory of Biomedical Engineering (currently Department of Biomedical Engineering and Computational Science) with the easiest job interview in my life so far. The former head of the Laboratory, Professor (emer.) Toivo Katila said: "You can go with Ilkka, he has something for you." Starting from that moment D.Sc. Ilkka Nissilä introduced me into the fascinating world of optical imaging. I would like to thank him for these years we have been working together.

Soon after that I started to build my own code for diffuse optical tomography. At that time, I was mentored by Professor Erkki Somersalo. I would like to thank him for all the knowledge he has passed to me in our face-to-face conversations. Also, his published materials have been excellent references.

In 2008, I made a short visit to University College London in UK. That was quite an exciting time. I would like to thank Professor Simon Arridge, who gave me the possibility to work in his research group, one of the leading groups in the field of optical imaging.

Most of the time, my doctoral studies have been unsupervised learning. In the beginning, I did not have an official supervisor for my thesis, but I had an opportunity to work with many great scientists. Only at the end, Professor Jouko Lampinen started to supervise my work. Even though he is not an expert in optical imaging, he was able to convert his knowledge

and expertise to the field. I only wish that we had started our collaboration earlier. I have been very pleased to work with him.

I would also like to thank everyone who has helped me preparing this thesis – PhD Tanja Tarvainen, D.Sc. Nuutti Hyvönen and the preliminary examiners PhD Ville Kolehmainen and PhD Oliver Dorn. It would not have been possible to complete this thesis without a supporting work environment. I would like to thank all the current and former members of the optical imaging group – Tiina Näsi, D.Sc. Jaakko Virtanen, Kalle Kotilahti, D.Sc. Juha Heiskala, D.Sc. Tommi Noponen, Lauri Lipiäinen, D.Sc. Jenni Heino, Atte Lajunen and many others. I also want to thank whole staff of the Department for all these years and memories that we have of the work and a little bit of the leisure as well.

Finally, I would like to thank my lovely better half Reetta, who has given me all her support during this work.

Espoo, 25 September 2011

Petri Hiltunen

Contents

List of Publications

This thesis consists of an overview and of the following publications which are referred to in the text by their Roman numerals.

I P. Hiltunen, D. Calvetti, and E. Somersalo. An adaptive smoothness regularization algorithm for optical tomography. *Optics Express*, 16(24), 19957–19977, 2008.

II J. Heiskala, P. Hiltunen and I. Nissilä. Significance of background optical properties and time-resolved information in diffuse optical imaging of term neonates. *Physics in Medicine and Biology*, 54(3), 535–554, 2009.

III P. Hiltunen, S.J.D. Prince, and S. Arridge. A combined reconstruction-classification method for diffuse optical tomography. *Physics in Medicine and Biology*, 54(21), 6457–6476, 2009.

IV P. Hiltunen, S. Särkkä, I. Nissilä, A. Lajunen, and J. Lampinen. State space regularization in the nonstationary inverse problem for diffuse optical tomography. *Inverse problems*, 27(2), 025009 (15 pages), 2011.

Author's Contribution

Publication I: "An adaptive smoothness regularization algorithm for optical tomography"

In this article, I was responsible for implementing the algorithm, designing the test cases and doing the computations. The article was written in collaboration with all three authors. The idea for the algorithm came from the third co-author.

Publication II: "Significance of background optical properties and time-resolved information in diffuse optical imaging of term neonates"

In this article, I was responsible for making the reconstructions based on the diffusion approximation. I implemented the finite element method for the forward problem and the optimization algorithm for the inverse problem. I participated in writing sections of the manuscript related to the diffusion approximation. The research problem was developed by the first and third co-authors.

Publication III: "A combined reconstruction-classification method for diffuse optical tomography"

In this article, I was responsible for implementing the algorithm, designing the test cases and carrying out the computations. I was mainly responsible for writing the article. The idea for method came from the third co-author and the algorithm development was shared work with all three authors.

Publication IV: "State space regularization in the nonstationary inverse problem for diffuse optical tomography"

In this article, I was responsible for implementing and running the nonstationary and stationary inverse problem algorithms. I designed the comparison with a previously presented method and how the reconstruction was carried from measured data. I was the main contributor of the article text except for the phantom and the measurement setup sections. The idea for the evolution model came from the second co-author and I developed the algorithm with the second co-author.

List of abbreviations

In alphabetic order

BEM	Boundary element method
BFGS	Broyden-Fletcher-Goldfarb-Shanno
BOLD	Blood oxygen level dependent
CFS	Cerebrospinal fluid
CM	Conditional mean
CT	Computed tomography
DA	Diffusion approximation
DOT	Diffuse optical tomography
EEG	Electroencephalography
EIT	Electrical impedance tomography
FDM	Finite difference method
FEM	Finite element method
fMRI	Functional magnetic resonance imaging
GPU	Graphics processing unit
iid	Identically independently distributed
MAP	Maximum a posteriori
MC	Monte Carlo
MEG	Magnetoencephalography
MRI	Magnetic resonance imaging
NIR	Near-infrared
RTE	Radiative transfer equation
PET	Positron emission tomography
pMC	Perturbation Monte Carlo
TV	Total variation

List of symbols

In order of appearance

$L(\mathbf{r}, t, \mathbf{s})$	radiance
$q(\mathbf{r}, t, \mathbf{s})$	source
c	speed of light
$\mathbf{r}$	spatial location
t	time
$\mathbf{s}$	scattering angle
μ_a	absorption coefficient
μ_s	scattering coefficient
$f_s(\mathbf{s}, \mathbf{s}')$	scattering phase function
Ω	domain
$\mathbb{R}$	real numbers
$u(\mathbf{r}, t)$	photon fluence
$q_0(\mathbf{r}, t)$	isotropic source
κ	diffusion coefficient
μ_s'	reduced scattering coefficient
g	anisotropy factor
ω	modulation frequency
$\mathbf{n}$	outer normal vector
R_n	reflection coefficient
n	refraction index
ξ	refraction index mismatch coefficient
θ_c	critical angle
$J^-(\mathbf{r}, t)$	injected inward directed boundary flux
T_n	transmission coefficient
Γ	exitance
$f(\cdot)$	forward operator
n_p	number of optical parameter

n_m	number of measurement
n_s	number of source
L^2	space of the square integrable functions, $\|f\|^2 = \int \|f\|^2 \mathrm{d}r < \infty$
x	optical parameters
y	measurement
σ_j	singular value
$\langle \cdot, \cdot \rangle$	inner product
$\dim(\cdot)$	dimension of the space
$\mathcal{R}(A)$	range of the linear operator A
$\mathcal{N}(A)$	null of the linear operator A
$X^\perp$	orthogonal complement of the space X
A^*	adjoint of the operator A
$A^\dagger$	pseudoinverse of the operator A
$\kappa(A)$	condition number of the operator A
$\\| \cdot \\|$	norm
α	regularization parameter
γ	step length parameter
$\partial\Omega$	boundary of the domain Ω
H^1	Hilbert space, $\|f\|^2 = \int \|f\|^2 \mathrm{d}r + \int \|f'\|^2 \mathrm{d}r < \infty$
$p(x)$	prior distribution
$p(y\|x)$	likelihood functional
$p(x\|y)$	posterior distribution
$p(y)$	prior predictive distribution

1. Introduction

1.1 Background

In medical imaging, the spectrum of the electromagnetic radiation is used broadly, such as x-rays and γ-rays in computed tomography (CT) and in positron emission tomography (PET), the radio frequencies in magnetic resonance imaging (MRI), and the visible light in endoscopes. A new emerging imaging modality called diffuse optical tomography (DOT) uses visible and near-infrared (NIR) light. Compared to methods using ionizing radiation (x-ray and γ-ray) the DOT method is non-ionizing and non-invasive. Typical building costs and the physical size of the device are smaller than in MRI or PET devices and no special environment, e.g. a magnetically shielded room, is needed. The instrumentation can be made portable and suitable for continuous bedside motoring, for monitoring infant and adult subjects. On the other hand the spatial resolution and depth sensitivity of DOT is lower than in MRI, but the temporal resolution can be higher [1].

The NIR light absorption in water at wavelengths 650 – 950 nm is low [2]. Therefore, NIR light can be used in biomedical applications, because the human tissue consists of mainly water and photons can propagate through the tissue without completely being absorbed. Propagation through tissue photons experience several elastic scattering events, where the photons do not lose energy but change their direction. A mathematical model for the photon propagation can be built on the absorption and the scattering coefficient of the tissue. The light propagation model is called the forward model, which gives the measured photon flux at the boundary of the illuminated medium when the optical properties of the medium, the absorption and the scattering coefficients, are known. In DOT, the photon flux is measured on the boundary. The process, in which the optical properties are determined based on given measurements is called the reconstruction of the optical properties and it is an inverse problem.

Medical imaging modalities can be divided into two different categories: structural and functional modalities. MRI and CT are examples of structural imaging modalities. They produce an image of the structures of the human body. Functional imaging modalities are, e.g., PET and functional MRI (fMRI). DOT can be seen as a structural imaging modality if different structures inside the tissue have different optical properties. Primarily, DOT is a functional imaging modality when spectroscopic imaging is done, i.e., several wavelengths are used. Extinction coefficients of oxygenated and deoxygenated haemoglobin at different wavelengths differ. Based on measurements of changes in absorption coefficients at several wavelengths, it is possible to make reconstructions of changes in de- and oxyhaemoglobin concentrations. In contrast to DOT, in fMRI the blood oxygen level dependent (BOLD) signal, which is sensitive to the paramagnetic properties of deoxyhaemoglobin, is measured.

DOT has potential for many interesting clinical applications in medical imaging because of its ability to reconstruct de- and oxyhaemoglobin concentration changes, tissue oxygen saturation and some other clinically interesting variables. Safe and portable imaging modality will provide a powerful tool for monitoring neonates. Hebden et al. [3] were first who made whole head images of the preterm infant who suffered from a cerebral haemorrhage within the left ventricle. They were able to see expected asymmetry in the blood volume images. Hebden et al. [4] have studied how the whole head optical properties change when the setting of the ventilation has changed in a preterm infant. Gibson et al. [5] have reconstructed changes in the haemoglobin concentrations induced from the motor evoked responses. There are also studies related to tomographic imaging of event related adult brain functions, such as [6, 7, 8, 9]. Another promising area would be breast cancer screening, e.g., [10, 11, 12, 13].

In multimodal imaging, several different imaging modalities are combined. DOT has great potential for multimodal imaging, because NIR light and optical fibers do not interfere with the magnetic or the electrical fields. Imaging devices such as magnetoencephalography (MEG) or electroencephalography (EEG) can be combined with DOT. This combination allows us to study simultaneously neuronal and heamodynamic activity. Combination of MRI and MEG is more technically challenging (cf. lowfield MEGMRI [14]).

1.2 Research objectives

The inverse problem in DOT is ill-posed, which mean that the problem does not have unique solution or the solution does not depend continuously on data, i.e. small errors in data can produce large errors in the solution. To overcome issues related to ill-posedness the Tikhonov regularization methods can be used. In this thesis, regularization methods for DOT are studied. Both stationary and nonstationary cases are considered. In this thesis the nonstationary inverse problem is defined such that an object is measured over a finite time period at constant time intervals and the optical properties of the object can change during the measurement. Then the optical properties at every time instance are to be reconstructed. In the stationary inverse problem the optical properties are static and one reconstruction is calculated.

The Tikhonov regularization methods have well known connection to the prior distributions in the statistical framework. In this thesis, the inverse problem is solved using classical Tikhonov regularization, but the connection is kept in mind in the design of the regularization functional. In this thesis, two different regularization methods for the static reconstruction problem in DOT was studied. They both allow discontinuities in the optical properties that might occur at boundaries between organs. For the nonstationary reconstruction problem, an efficient regularization model is presented.

The optical imaging device used in this thesis was constructed at Aalto University Department of Biomedical Engineering and Computational Science [15, 16]. It is a frequency domain system, i.e., the source light is intensity modulated. The attenuation of the amplitude and the phase shift of the light is measured. Therefore, in this thesis modelling of light propagation in the frequency domain is discussed. In DOT, two other widely used instrument types are the continuous wave and the time domain systems. Methods are tested in phantoms which are designed for certain fixed wavelengths only. Therefore, spectroscopic imaging is not discussed in this thesis.

2. Methods

In this section, the mathematical theory behind the forward model and the inverse problem is described. This section is divided into three subsections. In the section 2.1, the forward model which predicts the photon flux at the boundary given known optical parameters, is described. In this thesis, the diffusion approximation to the radiative transfer equation has been used. Theoretical background of the model is briefly discussed; a detailed derivation can be found in the references. Due to the approximations made, there are some limitations in the model that should be considered when the method is used in biomedical applications. In the end of the first subsection, a short review of different forward models used in DOT is given.

In the section 2.2, the theory behind the DOT inverse problem is discussed. First the basic theory of the linear inverse problems are reviewed. Since, the inverse problem of DOT is nonlinear, a nonlinear reconstruction method is presented. The inverse problem can be solved using the Bayesian framework, which is presented in this subsection.

In the section 2.3, nonstationary inverse problems are discussed. A state space model of the optical properties is presented in a statistical framework and the method of solving it, the Kalman filter, is presented. Several different regularization methods for this problem have been presented in the literature and these are discussed in this section.

2.1 Forward model

2.1.1 Physical model

Maxwell's equations describe the propagation of electromagnetic waves, such as light. However, for diffuse optical imaging, the wave theory can be ignored due to strong scattering [2]. Therefore, the radiative transfer equation (RTE), which is derived from the energy conservation of the particles, is a sufficient model for the propagation of photons. The RTE is an integro-differential equation

$$\frac{1}{c}\frac{\partial L(\mathbf{r},t,\mathbf{s})}{\partial t} + \mathbf{s}\cdot\nabla L(\mathbf{r},t,\mathbf{s}) + (\mu_a(\mathbf{r}) + \mu_s(\mathbf{r}))L(\mathbf{r},t,\mathbf{s}) = \mu_s(\mathbf{r})\int_{4\pi} f_s(\mathbf{s},\mathbf{s}')L(\mathbf{r},t,\mathbf{s})\mathrm{d}\mathbf{s}' + q(\mathbf{r},t,\mathbf{s}), \quad \mathbf{r}\in\Omega \tag{2.1}$$

where the radiance $L(\mathbf{r}, t, \mathrm{s})$ defines the power radiated through a given unit area and unit solid angle in a direction s at a location r and time t. The absorption coefficient $\mu_a(\mathbf{r})$ and the scattering coefficient $\mu_s(\mathbf{r})$ define the probabilities that a photon undergo an absorption or a scattering event in a unit length. The scattering phase function $f_s(\mathrm{s}, \mathrm{s}')$ is a probability density function, which defines the probability that a photon from a direction s scatters to the direction s$'$. The coefficient c is the speed of light in the medium $\Omega \subset \mathbb{R}^3$.

The RTE is computationally complex and sometimes it can be approximated by a simpler equation. In the P_N approximation, the radiance, the scattering phase function, and the source term are approximated by the truncated spherical harmonic series of order N. As a result we obtain a set of $(N+1)^2$ coupled partial differential equations. The scattering phase function is assumed to depend only on the cosine of the scattering angle, $f(\mathrm{s}, \mathrm{s}') = f(\mathrm{s}\cdot\mathrm{s}')$. An approximation, called the diffusion approximation (DA), which is frequently used in DOT, is derived from the P_1 approximation [17]. Many authors have presented the derivation of the P_N approximation, e.g. [18, 19]. DA approximates the photon fluence, also called the photon density, $u(\mathbf{r}, t) = \int L(\mathbf{r}, t, \mathrm{s})\mathrm{ds}$, by a partial differential equation

$$\frac{1}{c}\frac{\partial u(\mathbf{r},t)}{\partial t} - \nabla\cdot\kappa(\mathbf{r})\nabla u(\mathbf{r},t) + \mu_a(\mathbf{r})u(\mathbf{r},t) = q_0(\mathbf{r},t), \quad \mathbf{r}\in\Omega, \tag{2.2}$$

where $q_0(\mathbf{r}, t)$ is the isotropic component of the source

$$q_0(\mathbf{r},t) = \int q(\mathbf{r},t,\mathbf{s})\mathrm{d}\mathbf{s}. \tag{2.3}$$

The diffusion coefficient is $\kappa(\mathbf{r}) = (3(\mu_s'(\mathbf{r}) + \mu_a(\mathbf{r})))^{-1}$, where $\mu_s'(\mathbf{r}) = (1 - g)\mu_s(\mathbf{r})$ is the reduced scattering coefficient and the anisotropy factor g is the average cosine of the scattering angle

$$g = \int \mathbf{s}\cdot\mathbf{s}' f(\mathbf{s}\cdot\mathbf{s}')\mathrm{d}\mathbf{s}\mathrm{d}\mathbf{s}'. \tag{2.4}$$

In this thesis, a frequency domain system was used, where the source, $q_0(\mathbf{r}, t) = q_0(\mathbf{r})\exp(i\omega t)$, is modulated at an angular frequency ω. From which, the DA in the frequency domain is derived

$$-\nabla \cdot \kappa(\mathbf{r})\nabla u(\mathbf{r}) + \left(\frac{i\omega}{c} + \mu_a(\mathbf{r})\right) u(\mathbf{r}) = q_0(\mathbf{r}), \quad \mathbf{r} \in \Omega. \tag{2.5}$$

Validity of the DA in biomedical applications is discussed in section 2.1.3.

2.1.2 Boundary conditions, sources and measurables

If there are no photons entering the object, the boundary condition of the RTE is

$$L(\mathbf{r}, \mathbf{s}, t) = 0, \quad \mathbf{r} \in \partial\Omega,\ \mathbf{s} \cdot \mathbf{n} < 0, \tag{2.6}$$

where $\mathbf{n}$ is an outwards directed boundary normal vector. Because the DA does not contain any directional information, it cannot fully satisfy equation (2.6). Therefore approximations, such as the total photon flux entered into the object is zero,

$$\int_{\mathbf{s}\cdot\mathbf{n}<0} L(\mathbf{r}, \mathbf{s}, t)\mathbf{s} \cdot \mathbf{n}\mathrm{d}\mathbf{s} = 0, \tag{2.7}$$

have been presented by Ishimaru [20]. The presented boundary condition says that there is no inbound flux at the boundary and it does not take into account the reflections at the boundary. On the other hand, it is assumed that the reflected part of the outward directed flux is equal to the photon flux entering the object

$$\int_{\mathbf{s}\cdot\mathbf{n}<0} L(\mathbf{r}, \mathbf{s}, t)\mathbf{s} \cdot \mathbf{n}\mathrm{d}\mathbf{s} = \int_{\mathbf{s}\cdot\mathbf{n}>0} R_n(\mathbf{s})L(\mathbf{r}, \mathbf{s}, t)\mathbf{s} \cdot \mathbf{n}\mathrm{d}\mathbf{s}, \tag{2.8}$$

where $R_n(\mathbf{s})$ is the reflection coefficient which is dependent on the refraction index n of the object [21]. It is assumed that the refraction index of the object is constant and the object is surrounded by air ($n \approx 1$). Two different approximations of equation (2.8) have been presented in the literature: (i) $R_n(\mathbf{s})$ is approximated by two step functions [21] or (ii) $R_n(\mathbf{s})$ is approximated by a constant R_n, which is obtained experimentally [22]. Using the P_1 approximation of the radiance, both methods (i) and (ii) lead to the boundary condition in the frequency domain

$$u(\mathbf{r}) + 2\xi\kappa(\mathbf{r})\frac{\partial u(\mathbf{r})}{\partial \mathbf{n}} = 0, \quad \mathbf{r} \in \partial\Omega, \tag{2.9}$$

where the coefficient ξ takes into account mismatch of the refraction indices at the boundary. In the case (ii) the mismatch coefficient is $\xi = (1 + R_n)(1 - R_n)^{-1}$ [22] and in the case (i) the mismatch coefficient is

$$\xi = \frac{2(1 - R_0)^{-1} - 1 + |\cos\theta_c|^3}{1 - |\cos\theta_c|^2}, \tag{2.10}$$

where θ_c is the critical angle and $R_0 = (n-1)^2(n+1)^{-2}$ [21] . Both methods have been compared to each other by Schweiger et al. [23]. They found out that the difference between these two models is small in the time domain measurement model, where they measure the amplitude and the mean flight time of the photon package. The mean flight time and the phase shift are close to each other when the modulation frequency is less than 200 MHz [24]. The mismatch coefficient presented by Groenhuis et al. [22], $\xi = (1 + R_n)(1 - R_n)^{-1}$, was used in Publications I–IV.

Two different source conditions can be used: the collimated source model and the diffuse source model. In the collimated source model, a collimated photon beam at the boundary is modelled as an isotropic point source inside the object. The source function $q(\mathbf{r}) = Q\delta(\mathbf{r} - \mathbf{r}_q)$, where δ is the Dirac delta function, is located at the depth of the mean free path, $\mu_s'^{-1}$, below the boundary.

The diffuse source model can be presented as an inward directed boundary flux $J^-(\mathbf{r}, t)$, where the volume source term $q(\mathbf{r})$ is set to zero. The boundary condition 2.8 can be modified to

$$\int_{\mathbf{s}\cdot\mathbf{n}<0} L(\mathbf{r}, \mathbf{s}, t)\mathbf{s} \cdot \mathbf{n}\mathrm{d}\mathbf{s} = R_n \int_{\mathbf{s}\cdot\mathbf{n}>0} L(\mathbf{r}, \mathbf{s}, t)\mathbf{s} \cdot \mathbf{n}\mathrm{d}\mathbf{s} + (1 - R_n)J^-(\mathbf{r}, t), \quad (2.11)$$

where the transmission coefficient $T_n = 1 - R_n$ is calculated using the same approximation as in case (ii) above [25]. In the frequency domain, this leads to the boundary condition

$$u(\mathbf{r}) + 2\xi\kappa(\mathbf{r})\frac{\partial u(\mathbf{r})}{\partial \mathbf{n}} = 4J^-(\mathbf{r}), \quad \mathbf{r} \in \partial\Omega, \quad (2.12)$$

where the mismatch coefficient is $\xi = (1 + R)(1 - R)^{-1}$. A comparison of the source models is presented by Schweiger et al. [23]. Their results suggest that the models have a constant multiplicative difference except near sources and the collimated source condition is better matched with Monte Carlo simulation results. In Publications I–IV the diffuse source model was used, because the collimated source model requires information of the mean free path.

The measurable quantity is the photon flux coming out of the object, called the exitance, defined as [20]

$$\Gamma(\mathbf{r}, t) = \int_{4\pi} L(\mathbf{r}, \mathbf{s}, t)\mathbf{s} \cdot \mathbf{n}\mathrm{d}\mathbf{s}, \quad \mathbf{r} \in \partial\Omega. \quad (2.13)$$

Using the P_1 approximation, it gives

$$\Gamma(\mathbf{r}) = -\kappa\frac{\partial u(\mathbf{r})}{\partial \mathbf{n}} \quad (2.14)$$

in the frequency domain. Another measurement function has been presented

$$\Gamma_{\text{out}}(\mathbf{r}, t) = (1 - R) \int_{\mathbf{s}\cdot\mathbf{n}>0} L(\mathbf{r}, \mathbf{s}, t)\mathbf{s} \cdot \mathbf{n} ds \tag{2.15}$$

in [26, 27]. This will lead to exactly the same equation as equation (2.14) if the P_1 approximation is used.

Equations (2.5), (2.12), and (2.14) define the forward mapping[1], $f : \mathbb{R}^{n_p} \to \mathbb{R}^{n_m}$, where n_p is the number of the optical parameters and n_m is the number of the measurements, from the optical parameters to the measured flux at the boundary for given measurement locations $\{\mathbf{r}_{m,1}, \ldots, \mathbf{r}_{m,n_m}\}$ and source currents $J^{-}(\mathbf{r}_{s,i})$, at source locations $\{\mathbf{r}_{s,1}, \ldots, \mathbf{r}_{s,n_s}\}$, where n_s is the number of sources. There are several different methods, such as the finite difference method (FDM) and the finite element method (FEM), to numerically solve the elliptic partial differential equation defined by equations (2.5) and (2.12). In this case, FEM is more convenient because the domain Ω usually has a complex shape, especially in biomedical applications. Often, the optical parameters are presented in a small dimensional basis (dimension of n_p), which is interpolated to a more dense grid used in the FEM computation.

2.1.3 Validity of the physical model

The RTE itself is an approximation of the electromagnetic wave propagation. It does not take into account wave phenomena, such as diffraction, polarization, or changes in the refractive index [20, 18]. However, diffraction and polarization can be neglected in DOT due to strong scattering. The refraction index is mostly constant in the different tissues and the largest change in the refraction index occurs at the skin-air interface [2]. In the following, the approximations and assumptions behind the P_1 and DA are presented.

Firstly, the P_1 approximation is valid when $\mu_a \ll \mu_s$ [17]. In the biomedical application this assumption is mostly valid. One exception is the cerebrospinal fluid (CSF), which has a low scattering coefficient. The brain is surrounded by the CSF, which makes RTE more attractive for brain imaging, because it is valid in the low scattering regions. However, Custo et al. [28] presented, that the experimental reduced scattering coefficient

[1] Here we use discretized optical properties instead of optical parameters from space such as $L^2(\Omega)$. Similarly measurements are point measurements instead of continuous measurements over the boundary.

$\mu_s' = 0.3\,\mathrm{mm}^{-1}$ for small CSF regions in the adult studies, should not affect reconstructions much. However, there might be large regions of CSF in the the preterm infants' brain [Publication II]. Secondly, the P_1 approximation is not valid close to the source. The source-detector separation should be much larger than $1/\mu_s'$ [17]. This should be taken into account in designing the source-detector grid. However, this is not a major issue in difference imaging, where data is difference of the measurement at two different state, because some of the systematic errors are cancelled.

In addition to the limitations of the P_1 approximation, in DA it is assumed that the photon flux is slowly varying and the source is isotropic [17, 19]. In section 2.1.2 it was described how the beam from the source fibre can be modelled as an isotropic source. The DA assumes also that $c\mu_s'/\omega \gg 1$, i.e., the scattering frequency must be larger than the modulation frequency. This assumption is valid in biomedical applications if the modulation frequency is less than ~ 1 GHz [29].

2.1.4 Sensitivity of the measurement

The sensitivity of the measurement is defined as the change in the measured signal given a small change in the optical parameters. The sensitivity of the measurement defines the Jacobian matrix, $J_{i,j} = \partial\Gamma_i/\partial(\mu_a, \kappa)_j$, of the forward mapping. The sensitivity can be solved by perturbing every degree of freedom and calculating the corresponding change in the measurement [30]. This is called the direct form or the forward sensitivity analysis. It requires $n_s(n_p + 1)$ solutions of the forward equation. The computation time of the forward model is often long and the size of the basis, n_p, is large, therefore the forward sensitivity analysis is impractical.

Arridge [19] presented an efficient way to calculate the sensitivity of the measurement in DOT. This method is called the adjoint formulation or the adjoint sensitivity analysis. If the change in the diffusion coefficient at the boundary is assumed to be zero then the sensitivity of the measurement is

$$\delta\Gamma = -\int_\Omega \delta\kappa(\mathbf{r})\nabla v^*(\mathbf{r}) \cdot \nabla u(\mathbf{r}) + \delta\mu_a(\mathbf{r})v^*(\mathbf{r})u(\mathbf{r})\mathrm{d}\mathbf{r}, \tag{2.16}$$

where $\delta\mu_a(\mathbf{r})$ and $\delta\kappa(\mathbf{r})$ are changes in optical parameters and $v^*(\mathbf{r})$ is the solution of the adjoint equation

$$-\nabla \cdot \kappa(\mathbf{r})\nabla v^*(\mathbf{r}) + \left(-\frac{i\omega}{c} + \mu_a(\mathbf{r})\right) v^*(\mathbf{r}) = 0 \quad \mathbf{r} \in \Omega \tag{2.17}$$

$$v^*(\mathbf{r}) + 2\xi\kappa \frac{\partial v^*(\mathbf{r})}{\partial n} = q_m(\mathbf{r}) \quad \mathbf{r} \in \partial\Omega, \tag{2.18}$$

where q_m is a source at measurement location. Calculation of the Jacobian matrix requires only n_s solutions of the forward model and n_m solutions of the adjoint equation, which is more feasible method than the forward sensitivity analysis when $n_m < n_p$.

2.1.5 Other models used in the optical tomography

In this section, other methods and models which have been used in optical tomographic reconstructions, are briefly reviewed.

Regardless of the more complex computation of the RTE, it has been used as the forward model in the reconstructions in optical tomography by e.g. González-Rodríguez and Kim [31] and the FEM based solution of the RTE have been presented by Tarvainen et al. [32]. Also, stochastic modelling of light propagation using Monte Carlo (MC) simulation has been used. However, the MC based computation of the forward problem is the most computer time consuming and only reconstructions based on the linearized models are feasible, [33, 34, 35], because the computation of the Jacobian matrix is slow. Recent developments in graphics processing unit (GPU) computing has lead to remarkable speed-ups in the computation times, which make MC modeling more feasible for nonlinear reconstruction purposes [36, 37, 38].

There are other approximations of RTE than P_N, such as the Fokker-Planck equation [39]. This approximation is valid when scattering is forward peaked, i.e., $g \approx 1$. Most biological tissues are forward peaked (see table 1 of [2]). The Fokker-Planck equation approximates the photon propagation better than the DA in forward scattering media and low scattering areas [40]. González-Rodrígues and Kim [31] have made comparisons of RTE and Fokker-Planck based reconstructions, and the differences are not significant, but savings in the computation time were considerable.

Other than FEM based solutions of the DA have been presented in the literature, such as FDM and the boundary element method (BEM). FEM is more suitable than FDM for biomedical applications, because the FDM is practical to implement only in box shaped geometries and it does not allow denser grids near sources. BEM is suitable for cases when the object can be represented by a layered structure and the inverse problem is presented in terms of the location of the layers and the constant optical parameters within each layer. BEM based reconstructions are presented in [41, 42, 43]. Elisee et al. [44] presented a method which combines FEM and BEM, where the outer layers of the head (scalp, skull, brain) are discretized using BEM and the region inside the brain is modelled by FEM.

2.2 Inverse problem

2.2.1 Linear inverse problem

The inverse problem is defined as: given data y, find a variable x such that $y = f(x)$, where $f(\cdot)$ is a known linear or nonlinear mapping from the variable x to data y. A mathematical problem is well-posed if the solution fulfils three conditions stated by Hadamard: (i) the solution exists, (ii) the solution is unique, and (iii) the solution depends continuously on the data. Otherwise the problem is ill-posed. In this section, linear inverse problems are discussed, $f(x) = Fx$, where $F : X \to Y$ is compact, therefore continuous, and X and Y are finite or infinite dimensional Hilbert spaces. Every compact operator can be presented using the singular value decomposition

$$Fx = \sum_j \sigma_j \langle x, v_j \rangle u_j, \tag{2.19}$$

where σ_j are singular values, v_j and u_j are singular vectors and $\langle \cdot, \cdot \rangle$ is the inner product in space X. If $\dim(Y) = \infty$, then $\sigma_j \to 0$ when $j \to \infty$ [45].

Let us assume that data y belongs into the range of the operator F, $y \in \mathcal{R}(F)$, and the null space, or the kernel, of the operator F is trivial, $\mathcal{N}(F) = \{0\}$, then the system

$$y = Fx \tag{2.20}$$

has a unique solution $x = F^{-1}y$. The solution x is continuous on data if the operator F^{-1} is bounded [46]. Then the problem is well-posed, but given assumptions are difficult to fulfil in real applications. However, in case of the compact operators, the inverse operator is not bounded if X is not finite dimensional [47].

The data may not belong into the range of the operator F, if the data contains measurement noise or there is modelling errors. Then we can seek a solution which minimizes the norm of the residual $r = y - Fx$. Such a solution, the minimum residual solution, is an orthogonal projection of data to the range, $r \in R(F)^\perp$. The fundamental theorem of linear algebra states that $\mathcal{R}(F)$ is orthogonal to $\mathcal{N}(F^*)$, $\mathcal{R}(F)^\perp = \mathcal{N}(F^*)$ [47]. Therefore we can find a minimum residual solution using the equation $F^*(y - Fx) = 0$, which gives us the normal equation

$$F^*Fx = F^*y. \tag{2.21}$$

The solution of the normal equation, $x = (F^*F)^{-1}F^*y$, exists and is unique if and only if $\mathcal{N}(F^*F) = \{0\}$. Note that $\mathcal{N}(F) = \mathcal{N}(F^*F)$ [47]. The minimum residual solution is also known as the least squares solution. However, the solution does not depend continuously on data, because the operator F^*F is also compact [47].

Next we assume that the null space is not trivial. Given a vector $x_n \in \mathcal{N}(F)$, then all $x = x_r + x_n$, where $x_r \in \mathcal{N}(F)^\perp$ is a solution to equation (2.20), are solutions to equation (2.20). The generalized inverse or the pseudoinverse, $x = F^\dagger y$, defines the unique solution to the minimum residual problem, such that the solution has minimum norm, i.e. $\|x_n\| = 0$. If the measurement is inside the range $\mathcal{R}(F)$ and the null space of F is trivial, then pseudoinverse gives the classical solution $x = F^{-1}y$. If the measurement is not inside the range but the null space is trivial, then the pseudoinverse is the minimum residual solution. The pseudoinverse can be calculated using the singular value decomposition

$$F^\dagger y = \sum_j \frac{1}{\sigma_j} \langle y, u_j \rangle v_j, \tag{2.22}$$

where $\langle \cdot, \cdot \rangle$ is the inner product in space Y. See e.g. [48, 47, 45, 27].

The solution defined by the pseudoinverse fulfils the first two Hadamard conditions. However, it does not depend continuously on data in case of compact infinite dimensional operators[2], which means that the inverse problem is ill-posed [45]. Small changes in data, such as noise, can produce large changes in the solution.

In finite dimensional spaces, all operators are compact [46] and the matrix system $y = Fx$ has a unique pseudoinverse solution. In theory, the solution is continuously dependent on data [47]. In practice it might not be, due to numerical ill-conditioning of floating point arithmetic, i.e. the condition number $\kappa(F) = \|F\|\|F^\dagger\|$ is large. If the norm is the ℓ_2-norm, then the condition number is the ratio of the largest and the smallest singular value.

[2] $\|F^\dagger u_j\| = \sigma_j^{-1} \to \infty$, when $j \to \infty$

To overcome ill-posedness, Tikhonov suggested that the normal equation is altered to

$$(F^*F + \alpha I)x = F^*y, \tag{2.23}$$

which has the same solution as the minimization problem

$$x = \arg\min_{x} \|y - Fx\|^2 + \alpha\|x\|^2, \tag{2.24}$$

where $\alpha > 0$ is called the regularization parameter. Using the singular value decomposition, the solution can be presented as [27]

$$x = \sum_j \frac{\sigma_j}{\sigma_j^2 + \alpha} \langle y, u_j \rangle v_j. \tag{2.25}$$

The operator $F^*F + \alpha I$ is bounded from below, $\|(F^*F + \alpha I)x\| \geq \alpha\|x\|$, therefore the operator is invertible and continuous on data [47]. The Tikhonov regularized solution is the minimum residual solution with an augmented norm.

2.2.2 Nonlinear inverse problem in diffuse optical tomography

In the previous section, linear inverse problems were discussed. However, the inverse problem in DOT is nonlinear. The general theory of nonlinear inverse problems is far more complicated and is not discussed here in detail, but in can be found from reference such as [49, 48]. For example, uniqueness cannot be stated in case of the general nonlinear function. Even the minimum norm solution might not be unique. Next we assume that the Tikhonov regularized solution of the nonlinear problem is well-posed and postpone the discussion to section 2.2.3. In this section we discuss how to solve the given minimization problem numerically.

Let $f(x) : \mathbb{R}^{n_p} \to \mathbb{R}^{n_m}$ be the nonlinear finite-dimensional forward problem as defined in section 2.1.2. The Tikhonov regularized solution to the nonlinear inverse problem is presented as a minimization problem

$$x = \arg\min_{x \in \mathbb{R}^{n_p}} \|y - f(x)\|^2 + \alpha R(x) \tag{2.26}$$

similarly to the case of the linear problem. Here a generalized Tikhonov regularization is presented. A non-negative twice differentiable regularization functional $R(x)$ should reach its minimum in the proximity of the true solution. The selection of the regularization functional is discussed in section 2.2.6. Next we consider how the solution is found.

Traditional optimization algorithms can be used to minimize an objective functional

$$\Psi(x) = \|y - f(x)\|^2 + \alpha R(x). \tag{2.27}$$

Let assume that the functional $\Psi(x)$ is at least twice continuously differentiable. The quadratic approximation of the objective functional at x_0 is

$$q(\delta x) = \Psi(x_0 + \delta x) \approx \Psi(x_0) + \nabla\Psi(x_0)^{\mathrm{T}}\delta x + \frac{1}{2}\delta x^{\mathrm{T}} H_{\Psi}(x_0)\delta x, \qquad (2.28)$$

where $\nabla\Psi(x_0)$ and $H_{\Psi}(x_0)$ are the gradient and the Hessian matrix of the objective functional at x_0, respectively. A necessary condition of the local extreme of the quadratic approximation is $\nabla q(\delta x) = \nabla\Psi(x_0) + H_{\Psi}(x_0)\delta x = 0$. The direction $\delta x = -H_{\Psi}(x_0)^{-1}\nabla\Psi(x_0)$ is the descent direction of the objective functional, if the Hessian matrix is a symmetric positive definite matrix [50]. A minimum of the objective functional can be searched iteratively using an iteration

$$x_{k+1} = x_k - H_{\Psi}(x_k)^{-1}\nabla\Psi(x_k), \qquad (2.29)$$

which is known as Newton's method. In practise, the unit step might not ensure decrease in the objective functional. At every step, a variable step length $x_{k+1} = x_k + \gamma_k \delta x_k$ is taken. The step length is found from the optimization problem

$$\gamma_k = \arg\min_{\gamma \geq 0} \Psi(x_k + \gamma\delta x_k). \qquad (2.30)$$

Usually an inexact line search is used, such as the quadratic fit line search [50]. The line search ensures global convergence (if the Hessian matrix is not singular), i.e converge to a point where gradient is zero independently of the starting point. Another method to do that, even in the case of a singular Hessian matrix, is the Levenberg-Marquard (LM) method, where the Hessian matrix is replaced with a matrix $H_{\Psi}(x_k) + \tau_k I$, where an adaptive parameter $\tau_k > 0$ ensures positive definiteness of the replaced matrix.

In DOT, the Hessian matrix of the forward problem is complicated to compute. Methods which approximate the Hessian matrix, such as Gauss-Newton or quasi-Newton optimization methods, can be applied. An example of the quasi-Newton method is the Broyden-Fletcher-Goldfarb-Shanno (BFGS) method [50]. A comparison of the damped Gauss-Newton and the Levenberg-Marquard (with Gauss-Newton approximation of the Hessian matrix) methods in DOT is presented by Schweiger et al. [51]. They found out that the damped Gauss-Newton converged significantly faster. In Publications I–IV the damped Gauss-Newton method was used, where the Hessian matrix was approximated by $J(x_k)^{\mathrm{T}}J(x_k) + \alpha R''(x_k)$ and an inexact line search was utilized.

The optical parameters are positive physical quantities. This positivity constraint is implemented using the change of variables $\tilde{x} = \log x$ and the optimization is done using the variable $\tilde{x}$, which is unconstrained in $\mathbb{R}^{n_p}$, where as x lies in $\mathbb{R}_+^{n_p}$. Another possibility is to use constrained optimization, where the limitation $x > 0$ is implemented in the optimization algorithm. Suitable algorithms are, e.g., sequential quadratic programming, the gradient projection method and the trust-region method [50, 52].

2.2.3 Uniqueness, existence and continuous dependence of data in diffuse optical tomography

The general theory for nonlinear inverse problems can be found in, e.g., [49, 48]. There is presented sufficient conditions for existence and continuous dependence of data of the Tikhonov functional $\|y-f(x)\|^2+\alpha\|x-x_0\|^2$. Later in this section we present results related to DOT.

Uniqueness of DOT in the case of the DA is discussed in [53] and [54]. Arridge and Lionheart [53] showed that using intensity measurements only (steady state measurement, $\omega = 0$) it is not possible to separate the diffusion coefficient and the absorption coefficient. They also showed that if the refraction index, the diffusion coefficient, and the absorption coefficient are unknown, then there is no unique solution in both time and frequency domain measurements. However, if the refraction index is known then the diffusion coefficient and the absorption coefficient could be uniquely defined from the continuous measurement over the whole boundary. Later, Harrach [54] found out that, in case of a piecewise constant diffusion coefficient and a piecewise analytic absorption coefficient, the solution is unique in the steady state measurement, even in case of the complete measurement on small part of the boundary.

Egger and Schlottbom [55] investigated properties of the forward operator ($f(\kappa,\mu_a) : L^2(\Omega)\times L^2(\Omega)\to L^2(\partial\Omega)$) in DOT, such that could use the standard regularization theory of nonlinear functions. They showed basic properties of the forward operator such as continuity, differentiability, and compactness, i.e. the inverse problem is ill-posed as the solution does not depend continuously on the data. They showed that the forward operator fulfils criteria of the general theory of the nonlinear inverse problems (weakly closedness[3] and continuity) so that the Tikhonov functional

$$(\kappa, \mu_a) = \arg\min_{\kappa,\mu_a} \|y - f(\kappa, \mu_a)\|^2 + \alpha \left(\|\kappa - \kappa_0\|^2_{H^1(\Omega)} + \|\mu_a - \mu_{a,0}\|^2_{L^2(\Omega)} \right) \tag{2.31}$$

has a solution ($(\kappa, \mu_a) \in L^2(\Omega) \times L^2(\Omega)$), which depends continuously on data.

2.2.4 Selection of the regularization parameter

In the inverse problem literature, several methods to select the regularization parameter have been proposed, such as the Morozov discrepancy principle, the L-curve method, the generalized cross validation, and the unbiased predictive risk estimator [45, 48]. Most of these methods assume that the noise is identically independently distributed (iid) and that the level of the noise is known. These assumptions may, however, not hold true. However, the reconstructions in Publications II and IV implicitly assume iid. The nonlinear reconstruction in DOT requires a long computation time. Therefore, the regularization parameter selection procedures presented in the literature are usually impractical to implement. On the other hand, the value of the regularization parameter is of approximately the same magnitude in different applications and the reconstruction is relatively insensitive to its value. Therefore, selection of the regularization parameter was based on visual inspection of the results in Publications I–IV.

In statistical inversion methods, the regularization parameter selection can be seen as a careful analysis of the measurement noise and the available a priori information of the optical parameters (see section 2.2.5).

2.2.5 Bayesian framework

The solution of the inverse problem can be presented in the Bayesian framework, also known as statistical inversion theory. In the statistical framework, all unknown quantities are treated as random variables. Everything we know about the variable x before the measurement is presented as a distribution, known as the prior distribution $p(x)$. Similarly, the measurement model is presented as a conditional distribution $p(y|x)$, called the likelihood function. If the variable x is known, then all realizations of y (measurements) are from the distribution $p(y|x)$.

[3] with respect to weak topologies $H^1(\Omega) \times L^2(\Omega)$ and $L^2(\partial\Omega)$ [55]

Because all unknown quantities are random variables, the solution of the inverse problem is a distribution of these variables, called the posterior distribution. Using the Bayes' theorem, we can solve the posterior distribution of the variable x

$$p(x|y) = \frac{p(y|x)p(x)}{p(y)}, \tag{2.32}$$

where $p(y) = \int p(y|x)p(x)\mathrm{d}x$ is called the prior predictive distribution [56]. The posterior distribution reveals everything what we know about the variable x after the measurement with given a priori information and the assumed forward model. Only in very simple cases can the posterior distribution be solved in closed form. Therefore, two point estimates of the posterior density are used: the maximum a posteriori (MAP) estimate

$$x_{\mathrm{MAP}} = \max_x p(x|y), \tag{2.33}$$

and the conditional mean (CM) estimate

$$x_{\mathrm{CM}} = \int xp(x|y)\mathrm{d}x. \tag{2.34}$$

The CM estimate is optimal in the sense of the mean square error criterion whereas the MAP estimator penalizes errors uniformly [27].

The MAP estimate can be solved using minimization techniques described in section 2.2.2. In small dimensional problems, the integral in the CM estimate can be solved using numerical quadratures. However, in large dimensional problems the Monte Carlo sampling techniques have to be used [56]. The weakness of the sampling methods is that they are slow due to the slow computation of the forward model and large number of samples required. An advantage is that from sampled realizations of the x, one could calculate, e.g., the credibility interval of the CM estimate. An approximate credibility interval could be estimated using the normal approximation of the posterior distribution at the MAP estimate.

In section 2.2.4, it was mentioned that the regularization level is based on the subjective selection of the prior distribution and a careful analysis of the measurement noise. Let us assume that all distributions are Gaussian, $p(y|x) = N(y - f(x), \sigma_y^2 I)$ and $p(x) = N(x_0, \sigma_x^2 I)$ (for simplicity they are assumed to be iid). The negative logarithm of the posterior distribution is

$$-\ln p(x|y) = \frac{1}{2\sigma_y^2}\|y - f(x)\|^2 + \frac{1}{2\sigma_x^2}\|x - x_0\|^2 + \mathrm{Const.} \tag{2.35}$$

In this case, the MAP estimate corresponds to the Tikhonov regularized solution of the inverse problem, where the regularization parameter is defined as $\alpha = \sigma_y^2\sigma_x^{-2}$.

The prior distribution can contain an uncertain parameter θ, which is called a hyperparameter. Let us assume, that we know the prior distribution $p(x|\theta)$ for a fixed θ. As in the example above, one could select $\theta = \{x_0, \sigma_x\}$. The uncertainty in the parameter θ is presented by the distribution $p(\theta)$, called the hyperprior distribution. Model like this is called the hierarchical model. Given the likelihood function, one can show that the posterior density of the hierarchical model is

$$p(x, \theta|y) \propto p(y|x)p(x|\theta)p(\theta). \tag{2.36}$$

Hierarchical models have been used in Bayesian statistics, see, e.g., [56]. Similar, methods apply in the statistical inversion theory. Examples of the hierarchical models in inverse problems are presented, e.g, in Publication III and [57, 58, 59, 60]

2.2.6 Regularization of the inverse problem

In this section, different choices of the regularization functional $R(x)$ are discussed. As mentioned in section 2.2.5, there is a connection between the MAP estimate and Tikhonov regularization in the case of Gaussian distributions. This connection has been kept in mind when the regularization functionals were designed and analysed.

The simplest choice $R(x) = \|x\|^2$ will lead to Tikhonov regularization as presented in equation (2.24). In the Bayesian framework, this corresponds to the assumption that voxels x_i are iid. However, there is usually much more information which can be included in the regularization, such as the correlation between neighbouring voxels. The simplest framework to add this correlation, called smoothness regularization, is to use the regularization functional $R(x) = \|L(x - x_0)\|^2$, where the operator L is a differential operator, such as the Laplace operator (see e.g. [27, 48, 45])[4]. This corresponds to the improper Gaussian prior distribution where the mean is x_0 and the inverse of the covariance matrix is $L^{\mathrm{T}}L$. The distribution is improper because the matrix $L^{\mathrm{T}}L$ is singular. The method presented by Kaipio and Somersalo [27] or the one by Calvetti et al. [61] and a similar method presented in Publication I can be used to generate

[4]The smoothness regularization presented here is based on the norm $\int |\Delta x(\mathbf{r})|^2 \mathrm{d}\mathbf{r}$, i.e. the second order smoothness regularization. Others have used the norm $\int |\nabla x(\mathbf{r})|^2 \mathrm{d}\mathbf{r}$, i.e. the first order smoothness regularization, which leads to the discrete regularization term $x^{\mathrm{T}}Lx$ instead of $x^{\mathrm{T}}L^{\mathrm{T}}Lx$, where L is the discrete Laplace operator.

proper smoothness priors. Arridge et al. [62] have developed the smoothness prior further, such that there is a parameter which controls the correlation length of the covariance matrix. This type of prior density allows the inclusion of a priori information of the size of the objects inside the domain. Similarly, the prior density, which includes the correlation length, could be constructed using the Gaussian distribution with the covariance matrix from Matérn family of the covariance functions [63] as done in Publication IV.

Sometimes the smoothness assumption is not realistic. The parameter x might have discontinuities, while being otherwise smooth. One of the first regularization methods suitable for these kinds of objects was the total-variation regularization (TV) (see e.g. [27, 48, 45]). It has an edge preserving feature. The TV regularization functional

$$R(x) = \int_{\Omega} |\nabla x(\mathbf{r})| \mathrm{d}\mathbf{r} \tag{2.37}$$

for smooth enough functions is minimized when the length of the boundary of the inclusion multiplied with its height is minimized. Therefore, TV regularization favour smooth blocky objects. Examples of the application of TV regularization in DOT are given in[64, 65, 66].

TV regularization produces an unwanted staircase effect for smoothly varying parts of the image [67]. It has been also stated that in the case of very fine discretization, the CM estimate is not edge preserving in the statistical framework [68]. Douiri et al. [66, 69] presented an alternative for the edge preserving regularization, where discontinuities in the smoothness assumption are based on the gradient of the image. A similar approach has been presented in Publication I. The difference between these two methods is that the former is based on first order smoothness regularization and the latter is based on second order smoothness regularization. They have also different diffusivity functional.

In biomedical applications, other imaging modalities can be combined with DOT. The other modality, such as MRI, can be used as a priori information for DOT. It has been presented that the Gaussian prior distribution could be constructed from samples [27, 70]. As far as the author knows, this has not been used in DOT because samples from the optical parameter images are hard to obtain. Generally, it is assumed that structures in the optical and MR image corresponds to each other within reasonable accuracy. Several structural priors, which include, e.g., MRI information in the prior distribution for DOT have been presented, e.g. in [71, 72, 73, 74, 75, 76, 77].

2.3 Nonstationary inverse problem

2.3.1 State space model of the nonstationary inverse problem

In the nonstationary inverse problem, we have a time series of the measurements $\{y_1, y_2, \ldots, y_m\}$ called the measurement process, where the measurements are done at time instances t_k. Contrary to the stationary inverse problem, where the unknown variable x is assumed to be static, in the nonstationary inverse problem the variable x_k can vary over time.

As in section 2.2.5, here the measurement at t_k is a realization from a probability density function, $p(y_k|x_k)$, conditional on the state x_k. Similarly, propagation of the states is presented as a conditional probability distribution $p(x_{k+1}|x_k)$. It is assumed, that the process $\{x_1, x_2, \ldots, x_m\}$ is a Markov process, i.e. the next state x_{k+1} does not depend on the history of states $p(x_{k+1}|x_1, x_2, \ldots, x_k) = p(x_{k+1}|x_k)$, and it does not depend on previous measurements, i.e. $p(x_{k+1}|x_k, y_1, y_2, \ldots, y_k) = p(x_{k+1}|x_k)$.

For the given measurement process $\{y_1, y_2, \ldots, y_M\}$, the state process $\{x_1, x_2, \ldots, x_M\}$ can be estimated using a recursive method called optimal filtering or Bayesian filtering. The next state can be estimated from the evolution update equation, called Chapman-Kolmogorov equation,

$$p(x_{k+1}|y_1, \ldots, y_k) = \int p(x_{k+1}|x_k)p(x_k|y_1, \ldots, y_k)\mathrm{d}x_k, \tag{2.38}$$

where in the first step, the distribution $p(x_0|y_0)$ is equal to an initial prior distribution $p(x_0)$. When a new measurement y_{k+1} is observed, the posterior probability distribution can be calculated from the observation update equation

$$p(x_{k+1}|y_1, \ldots, y_{k+1}) = \frac{p(y_{k+1}|x_{k+1})p(x_{k+1}|y_1, \ldots, y_k)}{\int p(y_{k+1}|x_{k+1})p(x_{k+1}|y_1, \ldots, y_k)\mathrm{d}x_{k+1}}, \tag{2.39}$$

which is same as Bayes' theorem. See proof e.g. in [27, 78].

Let us assume that the measurement model is linear and the measurement noise is Gaussian, i.e.

$$y_k = F_k x_k + r_k, \quad r_k \sim N(0, R_k), \tag{2.40}$$

where R_k is the noise covariance. Similarly, we assume that the evolution model is linear and the error is Gaussian, i.e.

$$x_{k+1} = G_k x_k + q_k, \quad q_k \sim N(0, Q_k), \tag{2.41}$$

where Q_k is the model error covariance. Given these assumptions and using the evolution update equation and the observation update equation, one can show that the following recursion applies [27]. The evolution step is

$$\begin{aligned} m_k^- &= Gm_{k-1} \\ P_k^- &= GP_{k-1}G^{\mathrm{T}} + Q_{k-1}, \end{aligned} \tag{2.42}$$

and the estimation step is

$$\begin{aligned} S_k &= F_k P_k F_k^{\mathrm{T}} + R_k \\ K_k &= P_k^- F_k^{\mathrm{T}} S_k^{-1} \\ m_k &= m_k^- + K_k(y_k - F_k m_k^-) \\ P_k &= P_k^- - K_k S_k K_k^{\mathrm{T}}, \end{aligned} \tag{2.43}$$

where m_k and P_k are posterior CM estimates of the mean and the covariance and the matrix K_k is known as the Kalman gain. The method is the well-known Kalman filter, see e.g. [79, 80]. For nonlinear models, there are approximative methods, such as the extended Kalman filter and the unscented Kalman filter, see e.g. [79, 80, 81]. Like in statistical inverse problems, the posterior distribution can be explored using sampling methods. Similarly, there are sampling methods for Bayesian filtering, called particle filters, see e.g. [81].

In Publication IV, we used the Kalman filter for the linear approximation of the measurement model. The extended Kalman filter would have lead to long computation, because of the linearization of the forward model at every evolution step. An optimal smoother, see e.g. [82], would give a posterior density of every state given all measurements, $p(x_k|y_1,\ldots,y_M)$. This was not used because the basic Kalman smoother would require an inversion of a square matrix of the size of the state vector, which would take too long.

2.3.2 Regularization of the nonstationary inverse problems

A regularization method for the nonstationary inverse problem, based on the Tikhonov regularization, has been introduced for the ill-posed dynamical electric wire tomography [83, 84]. In the method, the Tikhonov functional $\|y_k - F_k x_k\|^2 + \alpha\|L(x_k - x_0)\|^2$ can be represented as an augmented measurement model

$$\|\tilde{y}_k - \tilde{F}_k x_k\|^2 = \left\| \begin{bmatrix} y_k \\ \sqrt{\alpha} L x_0 \end{bmatrix} - \begin{bmatrix} F_k \\ \sqrt{\alpha} L \end{bmatrix} x_k \right\|^2. \tag{2.44}$$

The augmented measurement model $\tilde{F}_k$ is replaced with the original forward model in the Kalman filter equations (2.42) and (2.43). The method has been used in the nonstationary electrical impedance tomography (EIT), e.g, [85, 86, 87, 88] and nonstationary DOT [65, 89]. One weakness of this method is that the size of the measurement model is large in three dimensional problems and the use of the Kalman filter is impractical. Later Kaipio and Somersalo [27] have presented that the augmented measurement model corresponds to an evolution model where the evolution operator G and the model error covariance are dependent of the matrix L in the regularization functional. This result suggests that the regularization can be achieved with an evolution model, where the covariance matrix has assumed temporal and spatial correlation structure. In Publication IV, the evolution of the states was modelled as an Ornstein-Uhlenbeck process [90] combined with a spatial covariance matrix from the Matérn family [63].

The augmented measurement model and the model presented in Publication IV imitate the smoothness regularization in the stationary inverse problems in both temporal and spatial domain. However, in some applications, such as in electrical impedance process tomography, the evolution of the states can be modelled accurately by equations modelling fluid flow. Then there is no need for extra regularization because a priori information of the states is exact. Examples of the process tomography can be found in, e.g., [91, 92, 93]. In the brain imaging, modelling the evolution of the states is much more complex (see [94]). Therefore, the Ornstein-Uhlenbeck process, which does not model physiology, but is the simplified model for the brain activity, was used.

On the other hand, states can be represented as a concatenated vector $[x_1, \ldots, x_M]^{\mathrm{T}}$, such that the process $\{y_1, \ldots y_M\}$ is

$$\begin{bmatrix} y_1 \\ \vdots \\ y_M \end{bmatrix} = \begin{bmatrix} F_1 & & 0 \\ & \ddots & \\ 0 & & F_M \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_M \end{bmatrix} + \begin{bmatrix} r_1 \\ \vdots \\ r_M \end{bmatrix}. \tag{2.45}$$

The Tikhonov regularized solution of this concatenated system is

$$\begin{bmatrix} x_1 \\ \vdots \\ x_M \end{bmatrix} = \arg\min_{x \in \mathbb{R}^{n_p M}} \left\| \begin{bmatrix} y_1 \\ \vdots \\ y_M \end{bmatrix} - \begin{bmatrix} F_1 & & 0 \\ & \ddots & \\ 0 & & F_M \end{bmatrix} x \right\|^2 + \alpha \|Lx\|^2. \tag{2.46}$$

If the regularization operator L has a block diagonal structure, the solution is the independent spatial Tikhonov regularized solution at every step. Non-diagonal blocks correspond to temporal regularization. An efficient method to solve the system (2.46) was suggested by Brooks et al. [95] in the nonstationary electrocardiography problem. Zhang et al. [96] have shown that this method corresponds to the Kalman filter and smoother with special choice of the regularization operator L. The concatenated system has been applied to spectroscopic DOT by Zhang et al. [97].

3. Summary and discussion of the publications

In this section results and findings of Publications I–IV are summarized.

3.1 Publication I: An adaptive smoothness regularization algorithm for optical tomography

In this study, an adaptive smoothness regularization scheme has been introduced for DOT. The correspondence between the regularization functional and Gaussian prior density was broadly used in the design of the smoothness regularization. The a priori assumption of the object was that it consists of regions of slowly varying optical properties and sharp variations at boundaries between regions. The discretization of the regularization functional was done in an unstructured grid. In complex domains, like in biomedical applications, generation of an unstructured grid is more practical than a structured one. One has to note that the unstructured grids in the forward model and in the regularization discretization were different.

The aristotelian approach (see [61]) was used to make the joint distribution of the variables on the boundary and the interior. Previously variance adjustment at the boundary has been presented for simple domains, such as an one dimensional line and a two dimensional rectangular grid [61]. In this paper, the adjustment was generalized to more complex domains. The coupling constant, which controls the mixing of the image regions, was updated during the iteration. The update of the regularization functional can be motivated by hierarchical Bayesian models.

The presented regularization method outperforms the regular smoothness regularization. The reconstructed images had higher contrast because discontinuities were allowed. However, one has to observe that if the object is not consistent with our a priori belief, then the regularization forces smooth transitions into sharp edges. It has been demonstrated in simpler applications that by calculating posterior statistics one can get better estimates of smooth transitions [98]. The computational time is increased eight-fold compared to the smoothness regularization due to the update of the regularization during the iteration.

Structural information can be incorporated into the regularization functional. The coupling constant which controls the diffusivity of the regularization functional was based only on the information in the image. This parameter could have information from an auxiliary structural image as demonstrated by Douiri et al. [69] and Arridge and Simmons [99].

3.2 Publication II: Significance of background optical properties and time-resolved information in diffuse optical imaging of term neonates

In this study, the effect of the optode grid geometry, time-resolved information and background optical properties were studied. The study was carried out using perturbation Monte Carlo (pMC) reconstruction technique. Monte Carlo simulation, which models light propagation by simulating individual photon interactions with matter, was used as the forward model, to generate simulated data and to generate Jacobian matrix. Both pMC and DA were used for reconstruction. Time-resolved information, such as the mean flight time of the photon and other statistics of the flight time, are collected using the time domain system. In the frequency domain, the mean flight time corresponds to the phase at the used modulation frequency [24].

Results suggest that a high density imaging geometry which includes both short and long source-detector separations, is optimal for functional brain activations studies in humans. The geometry presented by Zeff et al. [6] was less sensitive to the errors in the assumed background properties. Reconstructions were sensitive to the linearization point, i.e., assumed background properties, in sparse imaging geometries.

This study was mainly a feasibility study of the pMC method and the DA based methods were used for comparison. One has to note that DA based reconstruction methods has some limitations as discussed in the section 2.1.3. The object does not fulfil the condition $\mu_s' \gg \mu_a$ in CSF. It has been suggested that using an experimental value $\mu_s' = 0.3\ \mathrm{mm}^{-1}$

in the CSF instead of the correct value allows fairly correct results to be obtained with complex thin CSF regions [28]. This result is replicated by pMC in this paper. DA is not valid near the sources. In this study, the smallest source-detector separation was ~ 10 mm, which is much larger than ${\mu_s'}^{-1} \approx 0.5$ mm. On the other hand, data was difference data from a more accurate model. Difference imaging reduces systematic errors.

Background optical properties are essential in pMC method, which is based on linearization of the forward mapping. Accurate knowledge of the background optical properties is also relevant issue in the linearized DA model as used in Publication IV. The nonlinear reconstruction based on DA should not depend on the background optical properties because those can be reconstructed in principle.

3.3 Publication III: A combined reconstruction-classification method for diffuse optical tomography

In medical imaging, reconstruction and segmentation are often done successively. In this study, the combined reconstruction-classification method has been introduced for DOT. The main concept is to do reconstruction and segmentation/classification iteratively, so that the previous classification result is used in the regularization functional.

The presented method can be motivated by the Bayesian framework. Our a priori belief is that the optical parameter in every voxel is randomly drawn from the mixture of Gaussian distribution. The proposed method finds estimates for the optical parameters, the class probabilities and the mean and the variance of each class. Due to computational complexity, only the MAP estimate of the posterior distribution is used. The proposed method consists of iteration of reconstruction with the Tikhonov regularization functional and classification by the expectation-maximization algorithm.

The presented algorithm was tested on simulated data and a real phantom measurement. The prior model presented allows discontinuities in the optical parameters as in the model presented in Publication I. Therefore, much better contrast has been achieved than using traditional smoothness methods. Simultaneously method gives a probabilistic classification of the object. The classification error is smaller than in case of the successive smoothness regularization and classification steps.

The presented method quite flexibly transforms into different algorithms. If classes are known in advance, such as from a segmented anatomical image, then this method turns out to be the method presented by Guven et al. [72]. Heiskala et al. [35] proposed that a set of anatomical images could be used as a priori information in the image reconstruction. Similarly, the collected information could be used here as the prior distribution for class probabilities at every voxel.

3.4 Publication IV: State space regularization in the nonstationary inverse problem for diffuse optical tomography

In this publication, the regularization of the nonstationary inverse problems was discussed in contrast to Publication I and Publication III where the regularization of the stationary problem was discussed. Only a few imaging modalities, like DOT and EIT, need a regularization in the state-space model. Often regularization could be replaced by a quite accurate state evolution model. It has been suggested that Tikhonov regularization could be inserted as an augmented measurement. However, in this publication we presented a method which relies on the evolution model with a well chosen covariance structure.

The presented evolution model is a mean reverting process, the so called Ornstein-Uhlenbeck process. It is a stationary stochastic process, which will decay back to the average state if no measurements are made. At the stationary state, the spatial covariance is assumed to be from Matérn family of covariance functions. The Matérn family has the property that there are parameters which control the smoothness of the realizations and the correlation length.

For this paper also a phantom with dynamical features was developed. In that phantom, we are able to test the performance of the measurement device and the reconstruction methods. The phantom consists of a static body and two moving parts. The translating part simulates activations caused by the stimulus and the rotating part simulates noise caused by the background physiological activity.

Based on the test in simulated two dimensional geometry, the augmented model and presented model performs equally well regarding the image quality. The presented method has more degrees of freedom to control the smoothness and correlation length than the traditional smoothness regularization used in the augmented model. The advantage of the presented model is that the Kalman filter is computationally feasible in large discretizations, like in the three dimensional phantom study. It took about 45 minutes for the presented method to compute the reconstruction of the data. Based on an estimate, it would take 90 hours to compute the reconstruction using the augmented model. However, one has to note that the augmented model could be implemented as a smooth evolution process (see section 2.3.2).

In this method an auxiliary measurement, such as activation start times or the angular frequency of the noise process, are not used as inputs. If the noise process, which simulates background physiology, implemented as part of the phantom were stopped then the reconstructions were better. It would be useful to model the noise process to achieve better image quality. Diamond et al. [94] have done some work towards including the noise process into the reconstruction.

4. Discussion

In this thesis, we have developed several different regularization methods for stationary and nonstationary DOT. The classical Tikhonov regularization methods were used instead of the statistical inversion methods, however the connection was kept in mind during the design of the regularization functional. Much work was done on the regularization, but developing the measurement model, or in the statistical framework the likelihood functional, was not done here.

There have been much progress in the approximation error modelling, where errors are taken into account in the likelihood functional [62, 100, 101, 102]. Utilizing these methods more, one could make more realistic models of the errors in the measured signal. Error sources could be from, e.g., the measurement device, the approximative forward model, the incorrect or truncated shape of the object or the sparse discretization. A more accurate likelihood functional would make the Bayesian framework in DOT more feasible. Over or underestimated noise covariance would lead to too broad or too narrow posterior distribution. However, this does not affect on the value of the MAP estimate, but would make the statistical interpretation incorrect. Therefore, the Tikhonov-regularization method, which corresponds to the MAP estimate, was used in this thesis.

The feasibility of the developed regularization methods was shown using measurements from simple phantoms. However, in brain imaging for example, the object has a much more complex structure. Some additional information, such as anatomical information, would improve reconstructions. One could combine anatomical information in the regularizations presented in Publication I and Publication III as discussed in sections 3.1 and 3.3.

The DA based forward model was utilized in this thesis. A more accurate measurement model, based on RTE, would require much more computer time. Therefore, the DA is more feasible for nonlinear reconstructions. This might change in the near future. Both increasing computer power and recent developments in the GPU computation will make MC simulation of the photon propagation more reasonable than before.

The reconstruction methods for the nonstationary inverse problems in the DOT have been studied much less than reconstruction methods for stationary inverse problems even though many applications are nonstationary. In biomedical imaging, there is some interesting work, such as [97, 94], where auxiliary measurements and hemodynamic response function are included in the state space model. Fully 3D reconstruction of this kind would be good goal to pursue. In 3D, the degrees of freedom can easily explode to an impractical level.

Bibliography

[1] Huppert, T. J., Hoge, R. D., Diamond, S. G., Franceschini, M. A., and Boas, D. A. A temporal comparison of BOLD, ASL, and NIRS hemodynamic responses to motor stimuli in adult humans. *NeuroImage*, 29(2):368–382, 2006.

[2] Nissilä, I., Noponen, T., Heino, J., Kajava, T., and Katila, T. *Advances in Electromagnetic Fields in Living Systems*, volume 4, chapter Diffuse optical imaging, pages 77–129. Springer, 2005.

[3] Hebden, J. C., Gibson, A., Yusof, R. M., Everdell, N., Hillman, E. M. C., Delpy, D. T., Arridge, S. R., Austin, T., Meek, J. H., and Wyatt, J. S. Three-dimensional optical tomography of the premature infant brain. *Physics in Medicine and Biology*, 47(23):4155, 2002.

[4] Hebden, J. C., Gibson, A., Austin, T., Yusof, R. M., Everdell, N., Delpy, D. T., Arridge, S. R., Meek, J. H., and Wyatt, J. S. Imaging changes in blood volume and oxygenation in the newborn infant brain using three-dimensional optical tomography. *Physics in Medicine and Biology*, 49(7):1117, 2004.

[5] Gibson, A. P., Austin, T., Everdell, N. L., Schweiger, M., Arridge, S. R., Meek, J. H., Wyatt, J. S., Delpy, D. T., and Hebden, J. C. Three-dimensional whole-head optical tomography of passive motor evoked responses in the neonate. *NeuroImage*, 30(2):521–528, 2006.

[6] Zeff, B. W., White, B. R., Dehghani, H., Schlaggar, B. L., and Culver, J. P. Retinotopic mapping of adult human visual cortex with high-density diffuse optical tomography. *Proceedings of the National Academy of Sciences*, 104(29):12169, 2007.

[7] White, B. R., Snyder, A. Z., Cohen, A. L., Petersen, S. E., Raichle, M. E., Schlaggar, B. L., and Culver, J. P. Resting-state functional connectivity in the human brain revealed with diffuse optical tomography. *NeuroImage*, 47(1):148–156, 2009.

[8] Custo, A., Boas, D. A., Tsuzuki, D., Dan, I., Mesquita, R., Fischl, B., Grimson, W. E. L., and Wells, W. Anatomical atlas-guided diffuse optical tomography of brain activation. *NeuroImage*, 49(1):561–567, 2010.

[9] White, B. R. and Culver, J. P. Quantitative evaluation of high-density diffuse optical tomography: in vivo resolution and mapping performance. *Journal of Biomedical Optics*, 15:026006, 2010.

[10] Choe, R., Corlu, A., Lee, K., Durduran, T., Konecky, S. D., Grosicka-Koptyra, M., Arridge, S. R., Czerniecki, B. J., Fraker, D. L., DeMichele, A., Change, B., Rosen, M. A., and G., Y. A. Diffuse optical tomography of breast cancer during neoadjuvant chemotherapy: a case study with comparison to MRI. *Medical physics*, 32:1128–1139, 2005.

[11] van de Ven, S., Elias, S., Wiethoff, A., van der Voort, M., Leproux, A., Nielsen, T., Brendel, B., Bakker, L., van der Mark, M., Mali, W., and Luijten, P. Diffuse Optical Tomography of the Breast: Initial Validation in Benign Cysts. *Molecular Imaging and Biology*, 11(2):64–70, 2009.

[12] Nielsen, T., Brendel, B., Ziegler, R., Beek, M. v., Uhlemann, F., Bontus, C., and Koehler, T. Linear image reconstruction for a diffuse optical mammography system in a noncompressed geometry using scattering fluid. *Appl. Opt.*, 48(10):D1–D13, Apr 2009.

[13] Choe, R., Konecky, S. D., Corlu, A., Lee, K., Durduran, T., Busch, D. R., Pathak, S., Czerniecki, B. J., Tchou, J., Fraker, D. L., DeMichele, A., Change, B., Arridge, S. R., Schweiger, M., Culver, J. P., Schnall, M. D., Putt, M. E., Rosen, M. A., and Yodh, A. G. Differentiation of benign and malignant breast tumors by in-vivo three-dimensional parallel-plate diffuse optical tomography. *Journal of biomedical optics*, 14:024020, 2009.

[14] Magnelind, P. E., Gomez, J. J., Matlashov, A. N., Owens, T., Sandin, J. H., Volegov, P. L., and Espy, M. Co-registration of MEG and ULF MRI Using a 7 Channel Low-Tc SQUID System. *IEEE/CSC & ESAS European Superconductivity News Forum*, 4(14), 2010.

[15] Nissilä, I., Kotilahti, K., Fallström, K., and Katila, T. Instrumentation for the accurate measurement of phase and amplitude in optical tomography. *Review of scientific instruments*, 73:3306, 2002.

[16] Nissilä, I., Noponen, T., Kotilahti, K., Katila, T., Lipiäinen, L., Tarvainen, T., Schweiger, M., and Arridge, S. Instrumentation and calibration methods for the multichannel measurement of phase and amplitude in optical tomography. *Review of Scientific Instruments*, 76:044302, 2005.

[17] Boas, D. A. *Diffuse photon probes of structural and dynamical properties of turbid media: theory and biomedical applications*. Ph.D. thesis, University of Pennsylvania, 1996.

[18] Case, K. and Zweifel, P. *Linear transport theory*. Addison-Wesley, 1967.

[19] Arridge, S. R. Optical Tomography in Medical Imaging. *Inverse Problems*, 15(2):R41–R93, 1999.

[20] Ishimaru, A. *Wave propagation and scattering in random media*. Academic Press, 1978.

[21] Keijzer, M., Star, W. M., and Storchi, P. R. M. Optical diffusion in layered media. *Appl. Opt.*, 27(9):1820–1824, 1988.

[22] Groenhuis, R. A. J., Ferwerda, H. A., and Bosch, J. J. T. Scattering and absorption of turbid materials determined from reflection measurements. 1: Theory. *Appl. Opt.*, 22(16):2456–2462, 1983.

[23] Schweiger, M., Arridge, S. R., Hiraoka, M., and Delpy, D. T. The finite element method for the propagation of light in scattering media: Boundary and source conditions. *Medical Physics*, 22(11):1779–1792, 1995.

[24] Arridge, S. R., Cope, M., and Delpy, D. T. The theoretical basis for the determination of optical pathlengths in tissue: temporal and frequency analysis. *Physics in Medicine and Biology*, 37(7):1531, 1992.

[25] Tarvainen, T. *Computational Methods for Light Transport in Optical Tomography*. Ph.D. thesis, University of Kuopio, 2006.

[26] Heino, J. and Somersalo, E. Estimation of optical absorption in anisotropic background. *Inverse Problems*, 18(3):559, 2002.

[27] Kaipio, J. and Somersalo, E. *Statistical and Computational Inverse Problems*. Springer Verlag, 2004.

[28] Custo, A., III, W. M. W., Barnett, A. H., Hillman, E. M. C., and Boas, D. A. Effective scattering coefficient of the cerebral spinal fluid in adult head models for diffuse optical imaging. *Appl. Opt.*, 45(19):4747–4755, Jul 2006.

[29] Haskell, R. C., Svaasand, L. O., Tsay, T.-T., Feng, T.-C., McAdams, M. S., and Tromberg, B. J. Boundary conditions for the diffusion equation in radiative transfer. *J. Opt. Soc. Am. A*, 11(10):2727–2741, 1994.

[30] Arridge, S. R. and Schweiger, M. Photon-measurement density functions. Part 2: Finite-element-method calculations. *Appl. Opt.*, 34(34):8026–8037, 1995.

[31] González-Rodríguez, P. and Kim, A. D. Comparison of light scattering models for diffuse optical tomography. *Opt. Express*, 17(11):8756–8774, 2009.

[32] Tarvainen, T., Vauhkonen, M., and Arridge, S. R. Gauss–Newton reconstruction method for optical tomography using the finite element solution of the radiative transfer equation. *J. Qaunt. Spect. Rad. Transfer*, 109:2767–2278, 2008.

[33] Kumar, Y. P. and Vasu, R. M. Reconstruction of optical properties of low-scattering tissue using derivative estimated through perturbation Monte-Carlo method. *Journal of Biomedical Optics*, 9(5):1002–1012, 2004.

[34] Heiskala, J., Kotilahti, K., and Nissilä, I. An application of perturbation Monte Carlo in optical tomography. In *Engineering in Medicine and Biology Society, 2005. IEEE-EMBS 2005. 27th Annual International Conference of the*, pages 274–277. jan. 2005.

[35] Heiskala, J., Pollari, M., Metsäranta, M., Grant, P. E., and Nissilä, I. Probabilistic atlas can improve reconstruction from optical imaging of the neonatal brain. *Opt. Express*, 17(17):14977–14992, 2009.

[36] Alerstam, E., Lo, W. C. Y., Han, T. D., Rose, J., Andersson-Engels, S., and Lilge, L. Next-generation acceleration and code optimization for light transport in turbid media using GPUs. *Biomed. Opt. Express*, 1(2):658–675, Sep 2010.

[37] Fang, Q. and Boas, D. A. Monte Carlo simulation of photon migration in 3D turbid media accelerated by graphics processing units. *Opt. Express*, 17(22):20178–20190, Oct 2009.

[38] Ren, N., Liang, J., Qu, X., Li, J., Lu, B., and Tian, J. GPU-based Monte Carlo simulation for light propagation in complex heterogeneous tissues. *Opt. Express*, 18(7):6811–6823, Mar 2010.

[39] Arridge, S. R. and Schotland, J. C. Optical tomography: forward and inverse problems. *Inverse Problems*, 25(12):123010, 2009.

[40] Lehtikangas, O., Tarvainen, T., Kolehmainen, V., Pulkkinen, A., Arridge, S. R., and Kaipio, J. P. Finite element approximation of the Fokker-Planck equation for diffuse optical tomography. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 111(10):1406–1417, 2010.

[41] Sikora, J., Zacharopoulos, A., Douiri, A., Schweiger, M., Horesh, L., Arridge, S. R., and Ripoll, J. Diffuse photon propagation in multilayered geometries. *Physics in Medicine and Biology*, 51(3):497, 2006.

[42] Zacharopoulos, A. D., Arridge, S. R., Dorn, O., Kolehmainen, V., and Sikora, J. Three-dimensional reconstruction of shape and piecewise constant region values for optical tomography using spherical harmonic parametrization and a boundary element method. *Inverse Problems*, 22(5):1509, 2006.

[43] Zacharopoulos, A. D., Schweiger, M., Kolehmainen, V., and Arridge, S. 3D shape based reconstruction ofexperimental data in Diffuse OpticalTomography. *Opt. Express*, 17(21):18940–18956, Oct 2009.

[44] Elisee, J. P., Gibson, A., and Arridge, S. Combination of Boundary Element Method and Finite Element Method in Diffuse Optical Tomography. *Biomedical Engineering, IEEE Transactions on*, 57(11):2737–2745, nov. 2010.

[45] Vogel, C. R. *Computational methods for inverse problems*. Society for Industrial Mathematics, 2002.

[46] Kreyszig, E. *Introductory functional analysis with applications*. John Wiley & Sons New York, 1978.

[47] Lebedev, L. P., Vorovich, I. I., Gladwell, G. M. L., and ebrary, I. *Functional analysis: applications in mechanics and inverse problems*. Kluwer Academic Publishers, 2002.

[48] Engl, H. W., Hanke, M., and Neubauer, A. *Regularization of inverse problems*. Springer Netherlands, 1996.

[49] Engl, H. W., Kunisch, K., and Neubauer, A. Convergence rates for Tikhonov regularisation of non-linear ill-posed problems. *Inverse Problems*, 5(4):523, 1989.

[50] Bazaraa, M. S., Sherali, H. D., and Shetty, C. M. *Nonlinear programming: theory and algorithms*. John Wiley and Sons, 1993.

[51] Schweiger, M., Arridge, S. R., and Nissilä, I. Gauss-Newton method for image reconstruction in diffuse optical tomography. *Physics in medicine and biology*, 50(10):2365–86, 2005.

[52] Coleman, T. and Y., L. An Interior, Trust Region Approach for Nonlinear Minimization Subject to Bounds. *SIAM Journal on Optimization*, 6:418–445, 1996.

[53] Arridge, S. R. and Lionheart, W. R. B. Nonuniqueness in diffusion-based optical tomography. *Opt. Lett.*, 23(11):882–884, 1998.

[54] Harrach, B. On uniqueness in diffuse optical tomography. *Inverse Problems*, 25(5):055010, 2009.

[55] Egger, H. and Schlottbom, M. Analysis and Regularization of Problems in Diffuse Optical Tomography. *SIAM Journal on Mathematical Analysis*, 42(5):1934–1948, 2010.

[56] Gelman, A., Carlin, J. B., Stern, H. S., and Rubin, D. B. *Bayesian Data Analysis*. Chapman & Hall/CRC, 2004.

[57] Calvetti, D. and Somersalo, E. A Gaussian hypermodel to recover blocky objects. *Inverse Problems*, 23(2):733, 2007.

[58] Calvetti, D. and Somersalo, E. Hypermodels in the Bayesian imaging framework. *Inverse Problems*, 24(3):034013, 2008.

[59] Calvetti, D., Hakula, H., Pursiainen, S., and Somersalo, E. Conditionally Gaussian hypermodels for cerebral source localization. *SIAM Journal on Imaging Sciences*, 2(3):879–909, 2009.

[60] Bardsley, J. M., Calvetti, D., and Somersalo, E. Hierarchical regularization for edge-preserving reconstruction of PET images. *Inverse Problems*, 26(3):035010, 2010.

[61] Calvetti, D., Kaipio, J., and Somersalo, E. Aristotelian prior boundary conditions. *Int. J. Math. Comp. Sci*, 1:63–81, 2006.

[62] Arridge, S. R., Kaipio, J. P., Kolehmainen, V., Schweiger, M., Somersalo, E., Tarvainen, T., and Vauhkonen, M. Approximation errors and model reduction with an application in optical diffusion tomography. *Inverse Problems*, 22(1):175, 2006.

[63] Rasmussen, C. E. and Williams, C. K. I. *Gaussian Processes for Machine Learning (Adaptive Computation and Machine Learning)*. The MIT Press, 2005.

[64] Paulsen, K. D. and Jiang, H. Enhanced frequency-domain optical image reconstruction in tissues through total-variation minimization. *Appl. Opt.*, 35(19):3447–3458, Jul 1996.

[65] Kolehmainen, V., Prince, S., Arridge, S. R., and Kaipio, J. P. State-estimation approach to the nonstationary optical tomography problem. *J. Opt. Soc. Am. A*, 20(5):876–889, 2003.

[66] Douiri, A., Schweiger, M., Riley, J., and Arridge, S. Local diffusion regularization method for optical tomography reconstruction by using robust statistics. *Optics letters*, 30(18):2439–2441, 2005.

[67] Chan, T., Marquina, A., and Mulet, P. High-order total variation-based image restoration. *Siam J. ScI. comput*, 22(2):503–516, 2000.

[68] Lassas, M. and Siltanen, S. Can one use total variation prior for edge-preserving Bayesian inversion? *Inverse Problems*, 20(5):1537, 2004.

[69] Douiri, A., Schweiger, M., Riley, J., and Arridge, S. R. Anisotropic diffusion regularisation methods for diffuse optical tomography using edge prior information. *Meas. Sci. Tech.*, 18:87–95, 2006.

[70] Calvetti, D. and Somersalo, E. *Introduction to Bayesian scientific computing: ten lectures on subjective computing*. Springer, 2007.

[71] Intes, X., Maloux, C., Guven, M., Yazici, B., and Chance, B. Diffuse optical tomography with physiological and spatial a priori constraints. *Physics in Medicine and Biology*, 49(12):–155, 2004.

[72] Guven, M., Yazici, B., Intes, X., and Chance, B. Diffuse optical tomography with a priori anatomical information. *Phys. Med. Biol*, 50(12):2837–2858, 2005.

[73] Boverman, G., Miller, E. L., Li, A., Zhang, Q., Chaves, T., Brooks, D. H., and Boas, D. A. Quantitative spectroscopic diffuse optical tomography of the breast guided by imperfect a priori structural information. *Physics in Medicine and Biology*, 50(17):3941, 2005.

[74] Brooksby, B., Jiang, S., Dehghani, H., Pogue, B. W., Paulsen, K. D., Weaver, J., Kogel, C., and Poplack, S. P. Combining near-infrared tomography and magnetic resonance imaging to study in vivo breast tissue: implementation of a Laplacian-type regularization to incorporate magnetic resonance structure. *Journal of biomedical optics*, 10:051504, 2005.

[75] Yalavarthy, P. K., Pogue, B. W., Dehghani, H., Carpenter, C. M., Jiang, S., and Paulsen, K. D. Structural information within regularization matrices improves near infrared diffuse optical tomography. *Opt. Express*, 15(13):8043–8058, Jun 2007.

[76] Panagiotou, C., Somayajula, S., Gibson, A. P., Schweiger, M., Leahy, R. M., and Arridge, S. R. Information theoretic regularization in diffuse optical tomography. *J. Opt. Soc. Am. A*, 26(5):1277–1290, 2009.

[77] Fang, Q., Moore, R. H., Kopans, D. B., and Boas, D. A. Compositional-prior-guided image reconstruction algorithm for multi-modality imaging. *Biomed. Opt. Express*, 1(1):223–235, Aug 2010.

[78] Särkkä, S. *Recursive Bayesian Inference on Stochastic Differential Equations*. Ph.D. thesis, Helsinki University of Technology, 2006.

[79] Grewal, M. S. and Andrews, A. P. *Kalman filtering: theory and practice using MATLAB*. Wiley New York, 2001.

[80] Bar-Shalom, Y., Li, X. R., and Kirubarajan, T. *Estimation with applications to tracking and navigation*. Wiley-Interscience, 2001.

[81] Ristic, B., Arulampalam, S., and Gordon, N. *Beyond the Kalman filter: Particle filters for tracking applications*. Artech House Publishers, 2004.

[82] Brian, D., Anderson, O., and Moore, J. B. *Optimal Filtering*. Dower, 2005.

[83] Baroudi, D., Kaipio, J., and Somersalo, E. Dynamical electric wire tomography: a time series approach. *Inverse Problems*, 14(4):799, 1998.

[84] Kaipio, J. and Somersalo, E. Nonstationary inverse problems and state estimation. *journal of inverse and ill posed problems*, 7:273–282, 1999.

[85] Seppänen, A., Vauhkonen, M., Vauhkonen, P. J., Somersalo, E., and Kaipio, J. P. State estimation with fluid dynamical evolution models in process tomography - an application to impedance tomography. *Inverse Problems*, 17(3):467, 2001.

[86] Kolehmainen, V., Voutilainen, A., and Kaipio, J. P. Estimation of non-stationary region boundaries in EIT—state estimation approach. *Inverse Problems*, 17(6):1937, 2001.

[87] Vauhkonen, P. J., Vauhkonen, M., and Kaipio, J. P. Fixed-lag smoothing and state estimation in dynamic electrical impedance tomography. *International Journal for Numerical Methods in Engineering*, 50(9):2195–2209, 2001.

[88] Kim, B. S., Ijaz, U. Z., Kim, J. H., Kim, M. C., Kim, S., and Kim, K. Y. Non-stationary phase boundary estimation in electrical impedance tomography based on the interacting multiple model scheme. *Measurement Science and Technology*, 18(1):62, 2007.

[89] Prince, S., Kolehmainen, V., Kaipio, J. P., Franceschini, M. A., Boas, D., and Arridge, S. R. Time-series estimation of biological factors in optical diffusion tomography. *Physics in Medicine and Biology*, 48(11):1491–1504, 2003.

[90] Uhlenbeck, G. E. and Ornstein, L. S. On the Theory of the Brownian Motion. *Phys. Rev.*, 36(5):823–841, 1930.

[91] Seppänen, A., Heikkinen, L., Savolainen, T., Voutilainen, A., Somersalo, E., and Kaipio, J. P. An experimental evaluation of state estimation with fluid dynamical models in process tomography. *Chemical Engineering Journal*, 127(1-3):23–30, 2007.

[92] Seppänen, A., Vauhkonen, M., Vauhkonen, P. J., Voutilainen, A., and Kaipio, J. P. State estimation in process tomography—Three-dimensional impedance imaging of moving fluids. *International Journal for Numerical Methods in Engineering*, 73(11):1651–1670, 2008.

[93] Lipponen, A., Seppänen, A., and Kaipio, J. P. Reduced-order estimation of nonstationary flows with electrical impedance tomography. *Inverse Problems*, 26(7):074010, 2010.

[94] Diamond, S. G., Huppert, T. J., Kolehmainen, V., Franceschini, M. A., Kaipio, J. P., Arridge, S. R., and Boas, D. A. Dynamic physiological modeling for functional diffuse optical tomography. *Neuroimage*, 30(1):88–101, 2006.

[95] Brooks, D. H., Ahmad, G. F., MacLeod, R. S., and Maratos, G. M. Inverse electrocardiography by simultaneous imposition of multiple constraints. *Biomedical Engineering, IEEE Transactions on*, 46(1):3–18, 2002.

[96] Zhang, Y., Ghodrati, A., and Brooks, D. H. An analytical comparison of three spatio-temporal regularization methods for dynamic linear inverse problems in a common statistical framework. *Inverse Problems*, 21(1):357, 2005.

[97] Zhang, Y., Brooks, D. H., and Boas, D. A. A haemodynamic response function model in spatio-temporal diffuse optical tomography. *Physics in Medicine and Biology*, 50(19):4625, 2005.

[98] Calvetti, D. and Somersalo, E. A unified Bayesian framework for algorithms to recover blocky signals. In *Proceedings of SPIE*, volume 6697, page 669704. 2007.

[99] Arridge, S. R. and Simmons, A. Multi-spectral probabilistic diffusion using Bayesian classification. In B. ter Haar Romeny, L. Florack, J. Koenderink, and M. Viergever, editors, *Scale-Space Theory in Computer Vision*, volume 1252 of *Lecture Notes in Computer Science*, pages 224–235. Springer Berlin / Heidelberg, 1997.

[100] Kolehmainen, V., Schweiger, M., Nissilä, I., Tarvainen, T., Arridge, S. R., and Kaipio, J. P. Approximation errors and model reduction in three-dimensional diffuse optical tomography. *J. Opt. Soc. Am. A*, 26(10):2257–2268, Oct 2009.

[101] Tarvainen, T., Kolehmainen, V., Pulkkinen, A., Vauhkonen, M., Schweiger, M., Arridge, S. R., and Kaipio, J. P. An approximation error approach for compensating for modelling errors between the radiative transfer equation and the diffusion approximation in diffuse optical tomography. *Inverse Problems*, 26(1):015005, 2010.

[102] Tarvainen, T., Kolehmainen, V., Kaipio, J. P., and Arridge, S. R. Corrections to linear methods for diffuse optical tomography using approximation error modelling. *Biomed. Opt. Express*, 1(1):209–222, Aug 2010.

Aalto-DD 98/2011

Petri Hiltunen

Regularization methods for diffuse optical tomography

ISBN 978-952-60-4321-0 (pdf)
ISBN 978-952-60-4320-3
ISSN-L 1799-4934
ISSN 1799-4942 (pdf)
ISSN 1799-4934

Aalto University
School of Science
Department of Biomedical Engineering and Computational Science

BUSINESS +
ECONOMY

ART +
DESIGN +
ARCHITECTURE

SCIENCE +
TECHNOLOGY

CROSSOVER

DOCTORAL DISSERTATIONS

Aalto University