Dissertation

conducted at the

Institute of Medical Engineering

Graz University of Technology

Nonlinear reconstruction schemes and experimental design for fluorescence diffuse optical tomography

Manuel Freiberger

November 22th, 2010

Supervisor: ao.Univ.-Prof. Dipl.-Ing. Dr.techn. Hermann Scharfetter

Für Bianca

EIDESSTATTLICHE ERKLÄRUNG

Ich erkläre an Eides statt, dass ich die vorliegende Arbeit selbstständig verfasst, andere als die angegebenen Quellen/Hilfsmittel nicht benutzt und die den benutzten Quellen wörtlich und inhaltlich entnommenen Stellen als solche kenntlich gemacht habe.

Graz, am _____ _____

Datum Unterschrift

STATUTORY DECLARATION

I declare that I have authored this thesis independently, that I have not used other than the declared sources/resources and that I have explicitly marked all material which has been quoted either literally or by content from the used sources.

Graz, _____ _____

Date Signature

Abstract

Fluorescence diffuse optical tomography is a rather new optical imaging modality. The object, which is usually a small animal, is illuminated through a set of light sources placed on its surface. The light spreads inside the object and excites fluorophores which themselves emit photons at a longer wavelength. This secondary light is recorded by detectors on the boundary. From the knowledge of the light sources and the boundary measurements, it is sought to reconstruct the distribution of the fluorophore inside the object.

The reconstruction problem is nonlinear and ill-posed which makes it vulnerable to noise. In the first part of this thesis the use of nonlinear inversion schemes together with more advanced regularisation terms as total-variation regularisation and a method of levelset type will be investigated.

The second part deals with the optimal source and detector placement for a given target volume which has not been addressed so far. By minimising the redundancy in the measurements, the best optode locations can be selected and signals originating outside the target volume are suppressed.

Zusammenfassung

Die Fluoreszenzdiffusionstomographie ist ein optisches Bildgebungsverfahren mit dessen Hilfe Fluorophorverteilungen im Inneren eines Objekts wie zum Beispiel eines Kleintieres abgebildet werden können. Die Anregung des Fluorophors erfolgt durch Lichtquellen, die an der Objektoberfläche angebracht werden. Das Anregelicht breitet sich diffus im Körper aus, wird vom fluoreszierenden Stoff absorbiert, zum Teil als langwelligeres Licht wieder abgestrahlt und schließlich an der Oberfläche detektiert. Aus der Kenntnis der Lage der Quellen und der Intensität des Sekundärlichtes kann anschließend die Fluorophorverteilung rekonstruiert werden.

Dieses nichtlineare Rekonstruktionsproblem ist schlecht gestellt und wird von daher Messrauschen stark beeinflusst. Der erste Teil dieser Dissertation behandelt den Einsatz nichtlinearer Regularisierungsstrategien wie zum Beispiel “total-variation regularisation” zur Stabilisierung der Inversion.

Der zweite Teil befasst sich mit der optimalen Anordnung von Quell- und Empfangsoptoden für ein vorab definiertes Zielgebiet, ein Problem welches für Fluoreszenztomographie in dieser Art und Weise bisher noch nicht behandelt wurde. Die besten Optodenpositionen werden durch die Minimierung der Redundanz in den Messdaten bestimmt. Mit der erhaltenen Konfiguration lassen sich Fluoreszenzquellen außerhalb des Zielgebietes wirkungsvoll unterdrücken.

Contents

Nomenclature

BTE	Boltzmann's transport equation
CCD	Charge-coupled device
CW	continuous wave
FDOT	Fluorescence diffuse optical tomography
MC	Monte Carlo
PDE	Partial differential equation
SVD	Singular value decomposition
TV	Total-variation, page 28

C_{ad}	Admissible set of fluorophore concentrations, introduced in eq. (2.15), page 10
$H^1(\Omega)$	Space of functions over Ω with square-integrable gradients
$L^1(\Omega)$	Space of absolutely integrable functions over Ω
$L^2(\Omega)$	Space of square-integrable functions over Ω

α	Regularisation parameter, introduced in eq. (3.5), page 15
ϵ	Molar extinction coefficient, introduced in eq. (2.9), page 9
η	Fluorescence quantum yield, introduced in eq. (2.11), page 9
γ	Excitation-to-emission conversion factor, introduced in eq. (2.6), page 9
κ	Diffusion coefficient, introduced in eq. (2.3), page 8
μ_a	Absorption coefficient, introduced in eq. (2.1), page 5

μ_s	Scattering coefficient, introduced in eq. (2.1), page 5
μ_s'	Reduced scattering coefficient, page 8
ν	Speed of light in a medium, introduced in eq. (2.1), page 5
ϱ	Reflection coefficient, introduced in eq. (2.4), page 8
σ	Standard deviation, page 32
τ	Fluorescence life-time, introduced in eq. (2.11), page 9
φ	Photon density, introduced in eq. (2.3), page 8
ϕ	Levelset function, page 28
χ^2_k	Central chi-square distribution with k degrees of freedom
ψ	Adjoint photon density, page 20
ω	Modulation frequency of the light source, introduced in eq. (2.3), page 8
Θ	Scattering phase function, introduced in eq. (2.1), page 5
Φ	Photon density (direction dependent), introduced in eq. (2.1), page 5
Ω	Computation domain
$\partial\Omega$	Boundary of the computation domain

B	Finite-element mass matrix of boundary facets, introduced in eq. (2.18), page 11
K	Finite-element stiffness matrix, introduced in eq. (2.16), page 11
$\mathcal{M}$	Matrix of measurements, introduced in eq. (2.13), page 10
$\mathcal{M}^\delta$	Matrix of measurements corrupted by noise, page 14
M	Finite-element mass matrix, introduced in eq. (2.17), page 11
R	Matrix of residuals, introduced in eq. (4.15), page 32

---

cov	Covariance of a random variable
$\mathrm{cov}(a\|b)$	Conditional covariance of a given b
erf	Error function
F	Forward operator for fluorescence tomography, introduced in eq. (2.14), page 10

F'	Derivative of the forward operator; *sensitivity*, introduced in eq. (3.4), page 15
F'^*	Adjoint of the sensitivity, page 16
H	Entropy, introduced in eq. (6.9), page 49
H$(a\|b)$	Conditional entropy of a given b, introduced in eq. (6.11), page 49
H	Heaviside function, introduced in eq. (4.7), page 29
H_β	Relaxed Heaviside function, introduced in eq. (4.8), page 30
I	Identity operator
J	Derivative of the forward operator; *sensitivity*; see also F', page 45
MI	Mutual information, introduced in eq. (6.13), page 49
$\mathcal{R}$	Regularisation functional, introduced in eq. (3.5), page 15
Re	Operator for the real part
TV	Total-variation functional, introduced in eq. (4.4), page 28
var	Variance

$\mathbb{C}$	Field of complex numbers
CV	Coefficient of variation, page 41
c	Fluorophore concentration, introduced in eq. (2.9), page 9
d	Detector aperture function, page 9
g_T	Indices of nodes spanning the finite element T
$\mathcal{I}$	Set of measurement indices, page 45
N	Dimension of the computation domain (2D or 3D)
nc	Number of degrees of freedom of C_{ad}, page 19
nd	Number of photon detectors, page 10
$\mathcal{N}(\mu, \sigma^2)$	Normal distribution with mean μ and variance σ^2
ns	Number of light sources, page 10
nt	Number of mesh elements, introduced in eq. (6.6), page 47
q_{bc}	Boundary light source; used for modelling area illumination, introduced in eq. (2.2), page 5

q_{int}	Internal light source; mostly used for collimated beams, introduced in eq. (2.1), page 5
$\mathbb{R}$	Field of real numbers
S^{N-1}	Surface of the N-dimensional sphere, introduced in eq. (2.1), page 5

1 Introduction

> Aller Anfang ist leicht, und die letzten Stufen werden am schwersten und seltensten erstiegen.
>
> *(Johann Wolfgang von Goethe)*

The evolution of optical methods in biology and medicine was driven by the urge of man to understand themselves and their surrounding in tight embrace of technological possibilities. The investigations have been constituted from simple visual inspection of bodies with the unaided eye, the investigation of cavities by endoscopic means to the observation of small structures using microscopy and even tomographic imaging of cellular layers, by the invention of optical coherence tomography. While traditional methods are still in use due to their low cost and their ease of application, dedicated modalities haven been developed to overcome limits and insufficiencies in order to provide better image quality in terms of resolution or tissue contrast.

Fluorescence diffuse optical tomography (FDOT) is a rather new imaging technology. The investigated sample, which can be a slice of tissue or a small animal, for example, is illuminated through a set of light sources placed on the sample's surface. The excitation light spreads in the tissue where it is partially absorbed by fluorescent molecules, so called fluorophores. A part of the absorbed photons are re-emitted at a longer wavelength and propagate throughout the tissue again before they are recorded by detectors on the tissue boundary.

A particularly advantageous feature of fluorescence tomography is that the activity of fluorophores can be influenced by metabolic states or processes, e.g. the oxygenation of the tissue [1, 2], the pH value [3, 4] or the temperature [5]. This enables the opportunity of *functional imaging*, i.e., not only the anatomical structures can be visualised but also physiological states and activities in biological systems.

Nowadays, three different measurement principles for fluorescence tomography are in use which differ according to the type of excitation source and photon detectors in use: (i) continuous wave (CW) excitation [6], (ii) time-domain methods [7] and (iii) frequency-domain methods [8, 9]. CW systems employ light sources with constant intensity for the illumination of the sample. Usually, these systems are less expensive and also easier to implement compared to time- or frequency-domain methods. Unfortunately, these systems are not sensitive to time-varying processes and cannot monitor the fluorescence decay time, for example, but acquire only the steady-state, i.e. they fail to catch certain information from the inspected tissue which can be of relevance. This problem is circumvented by time-domain methods which measure the response of the tissue sample to a short excitation pulse and can thus be used for the determination of the fluorescence half-life. The drawback is the increased hardware cost for the generation and time-resolved measurement of short light pulses and the increased difficulty in handling. Frequency-domain setups are a compromise of CW and time-domain systems. Here the CW excitation intensity is modulated in a time-harmonic fashion and the amplitude and the phase are measured by the detectors. While the hardware is more complicated than in CW systems, e.g. due to the devices for intensity modulation, it is still less expensive than the equipment required for the generation of short laser pulses. From a signal-theoretic point of view the frequency-domain system is equivalent to time-domain measurements and the Fourier transform establishes the connection between both modalities. In this thesis the focus is on frequency-domain systems.

The spatial resolution of FDOT is rather low. The reason is that photons are scattered many times while traversing the tissue, and thus their trajectories can only be described in a statistical manner. This is also the origin for the ill-posedness of the inverse problem, i.e. the diffusive nature of light does not allow a reliable reconstruction of fine structures and the high spatial frequencies of the fluorophore concentration, respectively. In order to achieve a stable inversion, one has to use regularisation methods and add a-priori information during the reconstruction, which can be prior assumptions about the distribution of the fluorescent dye or the covariance of the measurement data, see e.g. [10].

For the reconstruction of the fluorophore distribution from measurements of the emitted light, most often linearisations of the forward problem are used in literature [11, 12]. When the non-linearity of the governing equations is considered,

usually quadratic regularisation terms like the L^2-norm are applied [13].

Not much work has been done regarding the optimal placement of optodes for fluorescence tomography. Up to now, the studies have focused on the influence of the number of sources and detectors and their distances on the singular values of the sensitivity matrix, which has to be computed in order to solve the inverse problem [14, 15]. However, these methods are only able to compare given measurement setups but do not provide means for the optimal placement of the optodes. Such a method is desired in practical applications where one is interested in imaging only a given volume of interest, e.g. a certain organ, and does not care about the surrounding.

This thesis seeks to partially overcome the previously mentioned limitations. The full nonlinear nature of the forward problem will be considered for solving the inverse problem. Additionally, non-quadratic regularisation methods, namely total-variation regularisation and a method of levelset type, are investigated to stabilise the inversion. By using relaxations of the true mathematical formulation, it is possible to devise implementations that are similar to ordinary L^2- and H^1-penalties and can therefore be easily incorporated into existing finite element codes.

For the adapted placement of optodes, a method originally used for the selection of the most independent measurement configuration in tomographic and microseismic monitoring is adapted to fluorescence tomography. This technique also allows one to define a target region prior to the selection of the optode positions and thus bias the measurement setup to this region. This is useful for an increased sensitivity in the target volume and to suppress off-focus signals as will be demonstrated later on.

The outline of this thesis is as follows: The mathematical description for light propagation in tissue is introduced in chapter 2. Then, chapter 3 deals with the reconstruction of fluorophore distributions from the knowledge of the light sources and the boundary measurements. In chapter 4 different quadratic and non-quadratic regularisation methods for the inverse problem are studied. The adaptive placement of optodes, given a volume of interest, is described in chapter 6. Finally, concluding remarks and an outlook can be found in chapter 7.

2 Forward model

> In der Schule haben wir gelernt, dass Lehrerwissen absolutes Wissen ist. Doch Wissenschaft kann niemals absolut sein. Sie ist die Kunst der Annäherung.
>
> *(Frank Schätzing)*

2.1 Models for light propagation

The forward model is the mathematical description of light propagation inside the tissue. In order to achieve physically correct models, it is necessary to incorporate the optical properties of tissue and light as well as the light-tissue interaction. Depending on the context, interesting properties of light are the intensity, the wavelength, the polarisation, the speed and the direction of propagation. The most important interaction processes with biological material are absorption and scattering. The first one describes the conversion of photon energy into other forms of energy, for example the excitation of electrons or vibrational energy of molecules, while the latter is the deflection of the photon out of its original trajectory into a new direction.

Monte Carlo methods: Monte Carlo (MC) methods for solving the forward model have a rather long tradition in optics. In the most simple form, they simulate, in every run, the path of a single photon throughout the tissue. As most of the physical processes like the change of the polarisation state or the absorption of photons by electron excitation are stochastic in their nature, MC methods are well suited to model these effects. Furthermore, it is comparatively easy to incorporate

non-homogeneous tissue and concomitant phenomena as reflection and refraction at internal boundaries due to varying indices of refraction.

The drawback of MC methods is their slow convergence which is proportional to the square root of the number of runs. In order to achieve reliable results for three dimensional objects, it is often necessary to simulate the propagation of hundreds of thousands or millions of photons. Furthermore, the convergence is poor when a certain physical process has a very low probability, say it occurs only once in one thousand trials. This has led to a series of strategies for speeding up the computation, for example, simulating photon packets rather than individual particles and randomly terminating photon packets that underwent too many absorption events [16]. Still, the relatively high computational effort limits the applicability of Monte Carlo methods for solving the inverse problem.

Transport equation: A deterministic solution for light propagation can be found in Boltzmann's transport equation (BTE) which is also frequently called the *radiative transport equation*. Rather than modelling each particle independently, which can be done with Monte Carlo methods, the BTE deals with an ensemble of photons in an infinitesimal volume. To be more precise, the BTE is a balance equation for the number of photons per unit volume (i.e. the photon density) $\Phi(x, s, t)$ at a spatial location x travelling into an infinitesimal solid angle around a direction s at a time instance t. The transport equation reads

$$\begin{aligned} \frac{1}{\nu}\frac{\partial \Phi(x,s,t)}{\partial t} + s \cdot \nabla \Phi(x,s,t) + (\mu_a(x) + \mu_s(x))\,\Phi(x,s,t) \\ = q_{int}(x,s,t) + \mu_s(x) \int_{S^{N-1}} \Theta(s,s')\Phi(x,s',t)\,ds' \qquad \text{in } \Omega, \end{aligned} \tag{2.1}$$

together with the boundary condition

$$\Phi(x,s,t) = q_{bc}(x,s,t) \qquad \forall s : s \cdot n < 0 \qquad \text{on } \partial\Omega, \tag{2.2}$$

where ν is the speed of light in the tissue, μ_a and μ_s are the absorption and scattering coefficients, respectively, and $\Theta(s, s')$ is the scattering phase function which states the probability that a photon travelling in direction s' will be deflected into direction s. In this equation, S^{N-1} is the surface of the N-dimensional sphere and thus the

integral is taken over all possible directions. The term q_{int} models sources inside the domain.

The left-hand side of equation (2.1) models the decrease in photon density due to movement and absorption, while the right-hand side incorporates the photon gain caused by scattering and an internal source.

The boundary condition imposes restrictions on the solution only for inward-pointing directions, i.e. for all s with $s \cdot n < 0$ where n is the outward-pointing unit normal of the domain. It states that no light can enter the domain from the outside except at locations where a boundary source is present, i.e. where $q_{bc}(x, s, t) \neq 0$.

A limitation of the transport equation is that it cannot handle wave effects. As a consequence, the wavelength of the light must be much smaller than the dimensions of the structures in the domain [17]. Furthermore, varying indices of refraction are not included in equation (2.1) but can be incorporated.

Despite these insufficiencies, the transport equation would be well suited for biomedical optics. It is able to handle many different types of tissue whether they are highly scattering like muscle, as absorbing as the liver or as transparent as liquor, for example. Furthermore, the boundary equation is physically meaningful.

The reason why the BTE is frequently avoided is the high computational effort required in order to achieve an accurate solution. The photon density $\Phi(x, s, t)$ has to be discretised in the spatial and angular domain which requires five variables for a 3D problem (three coordinates and two angles). If too little angular directions are chosen for the discretisation, the resultant photon field favours propagation along these directions resulting in a star-like pattern as can be seen in 2.1 where one would expect a radial symmetric distribution instead. To illustrate the influence of the angular discretisation on the computational effort, imagine a rather coarse discretisation of only 32 divisions for both the azimuth and the inclination. This results already in 1024 directions which has to be multiplied by the number of mesh points of the spatial discretisation.

Diffusion approximation: Due to the shortcomings of the previously described models, the most common description for photon migration is the diffusion approximation. This is a reduced form of the BTE which first expands the direction-

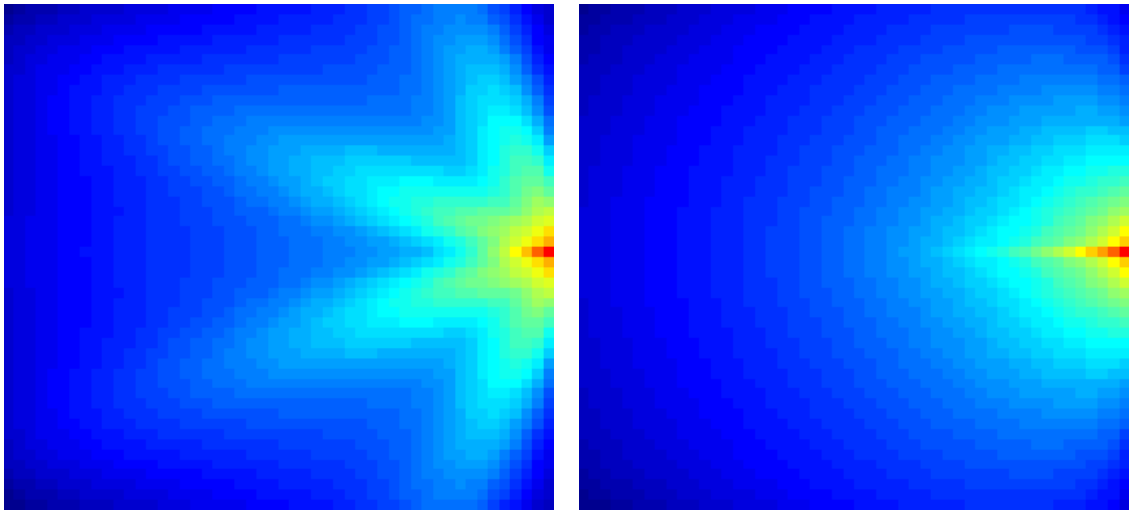

Figure 2.1: Photon distribution (logarithmic) for a homogeneous medium with absorption $\mu_a = 0.01\,\mathrm{mm}^{-1}$ and scattering $\mu_s = 5\,\mathrm{mm}^{-1}$. The angular domain was discretised with eight directions (left) and 24 directions (right). The star-like light propagation when using too little directions is clearly visible in the left image. The images are by courtesy of Matthias Schlottbom.

dependent quantities Φ, Θ, q_{int} and q_{bc} into spherical harmonics. Then, all spherical harmonics except the first one are neglected, i.e. only the isotropic propagation of photons is considered. Some additional constraints are imposed on the photon flux and the scattering phase function Θ. A thorough derivation is found in the review article [18]. The model is valid in the regime where photon propagation is diffusive, which is the case in tissues which are much more scattering than absorbing, at a sufficient distance from the source. Problems arise if the tissue is either strongly absorbing such as in the liver or nearly transparent as in liquor. Furthermore, the model is inaccurate near the light source, where the photon propagation is not yet diffusive due to the small number of scattering events, and at places where the optical properties jump which might be the case at the boundary of organs, for example [17].

In its time-harmonic form, the diffusion approximation is a partial differential equa-

tion (PDE) which reads

$$-\nabla \cdot (\kappa(x)\nabla\varphi(x)) + \left(\mu_a(x) + \frac{i\omega}{\nu}\right)\varphi(x) = q_{int}(x), \qquad \text{in } \Omega, \tag{2.3}$$

$$\varrho\varphi(x) + \kappa(x)n \cdot \nabla\varphi(x) = \begin{cases} \varrho\, q_{bc}(x), & \text{on } \Gamma_{bc}, \\ 0, & \text{on } \partial\Omega \setminus \Gamma_{bc}, \end{cases} \tag{2.4}$$

where φ is the photon density, $\kappa = (N(\mu_s' + \mu_a))^{-1}$ the diffusion coefficient of the tissue, μ_s' the reduced scattering coefficient and μ_a the absorption coefficient. ω is the modulation frequency of the excitation light source with the special case of $\omega = 0\,\mathrm{s}^{-1}$ for continuous wave excitation with a constant intensity. The light injected can be modelled by an internal source q_{int} which is frequently used for collimated beams or a boundary source q_{bc} suitable for spatially extended sources on a part Γ_{bc} of the boundary $\partial\Omega$. ϱ is a reflection coefficient for the incorporation of reflections at the boundary $\partial\Omega$ of the domain Ω due to changes in the index of reflection.

2.2 Diffusion approximation for fluorescence tomography

A straightforward way to apply the diffusion approximation to the field of fluorescence tomography is to use a two state model [13, 19]: one equation describes the propagation of the excitation light which is injected at the surface of the sample and another one is used for the propagation of light emitted by the fluorophore inside the object. The two equations are coupled by a term that converts light from the excitation wavelength to the emission wavelength. In reality, a fluorophore exhibits a band-limited continuous emission spectrum leading to a continuous distribution of the fluorescence wavelength. In favour of simplicity, the continuous spectra are neglected and only two discrete wavelengths are considered.

The forward model for fluorescent tomography consists of two coupled PDEs and has the form

$$-\nabla \cdot (\kappa_x \nabla\varphi_x) + \mu_x\varphi_x = q_{int}, \qquad \text{in } \Omega, \tag{2.5}$$

$$-\nabla \cdot (\kappa_m \nabla\varphi_m) + \mu_m\varphi_m = \gamma c\varphi_x, \qquad \text{in } \Omega, \tag{2.6}$$

with the boundary conditions

$$\varrho_x \varphi_x + \kappa_x n \cdot \nabla \varphi_x = \begin{cases} \varrho_x q_{bc}, & \text{on } \Gamma_{bc}, \\ 0, & \text{on } \partial\Omega \setminus \Gamma_{bc}, \end{cases} \tag{2.7}$$

$$\varrho_m \varphi_m + \kappa_m n \cdot \nabla \varphi_m = 0, \qquad \text{on } \partial\Omega, \tag{2.8}$$

where φ_i, $i \in \{x, m\}$ denote the complex amplitudes of the excitation (x) and the emitted (m) light. If it is needed to emphasise the dependence of the photon densities on the source, the notation $\varphi_i(q)$ will be used.

The diffusion and (complex) absorption coefficients have to be extended to incorporate the dependence on the fluorophore concentration c. They are defined by

$$\kappa_i(x) = \frac{1}{N\left(\mu_{a,i}(x) + \mu'_{s,x}(x) + \epsilon_i c(x)\right)}, \tag{2.9}$$

$$\mu_i(x) = \mu_{a,i}(x) + \epsilon_i c(x) + \frac{i\omega}{\nu}, \tag{2.10}$$

with $i \in \{x, m\}$. In these equations, $\mu_{a,i}$ serves as background absorption which is increased by the product $\epsilon_i c$ when a fluorophore is present. ϵ_i is the molar extinction coefficient of the fluorophore at the respective excitation or emission wavelength. This quantity links the fluorophore's concentration and its absorption coefficient.

The scalar conversion factor on the right-hand side of (2.6) is defined as

$$\gamma = \frac{\eta}{1 - i\omega\tau} \epsilon_x. \tag{2.11}$$

Here, η is the quantum yield of the fluorophore (i.e. the ratio of the absorbed photons to the ones emitted), and τ is the fluorescence life-time. According to (2.6), many photons will be emitted at locations where sufficient absorption is present at the excitation wavelength (the product $\epsilon_x c(x)$) and where the conversion efficiency η is high. It is straightforward to generalise this quantity by the assumption that the fluorescence life-time and/or the quantum yield are spatially variant, i.e. $\tau \to \tau(x)$ and $\eta \to \eta(x)$.

A measurement in fluorescence tomography is defined as the flux of the photons leaving the domain at a detector site. The function $d(s)$, $s \in \partial\Omega$, incorporates the detector's aperture and its transfer characteristics. In the following, it is assumed

to be real-valued but it can be extended to a complex-valued function without problems. With this, the definition of the measurement can be written as

$$-\int_{\partial\Omega} \kappa_m n \cdot \nabla\varphi_m d\, ds \overset{(2.8)}{=} \varrho_m \int_{\partial\Omega} \varphi_m d\, ds. \tag{2.12}$$

For a tomographic measurement process, the object under investigation is illuminated with ns light sources from different angles and all nd detectors are read out simultaneously. The measurements can be arranged in a matrix $\mathcal{M}$ where the ij-th entry is the reading of the i-th detector when the sample is illuminated by the j-th source, i.e.

$$\mathcal{M}_{ij} := \varrho_m \int_{\partial\Omega} \varphi_m(q_j) d_i\, ds. \tag{2.13}$$

Finally, we can introduce the forward operator F that relates a concentration distribution c to the matrix of measurements $\mathcal{M}$, i.e.

$$\begin{aligned} F:\ C_{ad} \subset L^2(\Omega) &\to \mathbb{C}^{nd\times ns} \\ c &\mapsto \mathcal{M} := [\mathcal{M}_{ij}]_{1\le i\le nd, 1\le j\le ns}, \end{aligned} \tag{2.14}$$

where the admissible set C_{ad} of fluorophore is defined as

$$C_{ad} := \{c \in L^2(\Omega) : 0 \le c \le \overline{c} \quad \text{in } \Omega\}, \tag{2.15}$$

i.e. the upper concentration is bounded by $\overline{c} > 0$ which in turn implies that $\kappa_i > 0$ and $\mu_i < \infty$. For a mathematical discussion on the properties of the forward operator see [20] and the references therein. We use the notation $F_{ij}(c)$ to denote the mapping of the fluorophore concentration c to the i-th detector reading when the sample is excited by the j-th source.

2.3 Discretisation

There are various possibilities for the discretisation of the diffusion approximation. Two of those which have been applied to biomedical optics are the finite difference method [9] and the finite element method [13]. In this thesis, the focus is put

on the latter. The reason is that finite element methods allow modelling complex anatomical geometries. Also, different boundary conditions of Dirichlet, Neumann or Robin type can be treated in a consistent manner.

A standard finite element method [21] with tetrahedral elements is chosen for the discretisation of the system (2.5)–(2.8) of partial differential equations. The photon fields φ_i, the associated sources q_{int} and q_{bc} and the detectors d are discretised with piecewise linear basis functions. Also, the different optical parameters κ_i, μ_i as well as the fluorophore concentration c are expanded in the same set of basis functions.

In the following, the construction of the finite element system shall be explained in short. Let g_T be the indices of the four nodes spanning the tetrahedron T and ψ the linear basis functions on the finite element mesh. Accordingly, g_F shall be the three node indices spanning a face F of a tetrahedron on the outer boundary $\partial\Omega$ of the computation domain Ω.

The 4×4 contributions of the tetrahedron T to the stiffness and the mass matrix are then given by

$$K_{ij}^T(\kappa) := \int_T \kappa \nabla \psi_{g_T(i)} \cdot \nabla \psi_{g_T(j)} dx, \qquad \text{and} \tag{2.16}$$

$$M_{ij}^T(\mu) := \int_T \mu \, \psi_{g_T(i)} \psi_{g_T(j)} dx \qquad 1 \leq i, j \leq 4. \tag{2.17}$$

The result is similar for the 3×3 mass matrix of the boundary term which is

$$B_{ij}^F(\varrho) := \varrho \int_F \psi_{g_F(i)} \psi_{g_F(j)} dx \qquad 1 \leq i, j \leq 3. \tag{2.18}$$

The assembled finite element system is of the form

$$[K(\kappa_x) + M(\mu_x) + B(\varrho_x)]\varphi_x = q, \tag{2.19}$$

$$[K(\kappa_m) + M(\mu_m) + B(\varrho_m)]\varphi_m = M(\gamma c)\varphi_x. \tag{2.20}$$

In this system K, M and B are the assembled stiffness, mass and boundary mass matrices, v_x and v_m are the photon fields at the excitation and emission wavelength and q is the source term including both the internal and the boundary source. The

measurement $\mathcal{M}_{ij}$ of the i-th detector d_i and the j-th source q_j is just the inner product

$$\mathcal{M}_{ij} := d_i^\top v_m(q_j). \tag{2.21}$$

2.4 Examples

A showcase result of the photon density distribution in highly scattering material is given in figure 2.2. The photon fields were simulated on a disc with a diameter of 30 mm. The optical properties assigned to the phantom have been compiled from literature [22, 23, 24, 25]. Their values are listed in table 2.1. The modulation frequency is set to $\omega = 0\,\mathrm{s}^{-1}$ which is—mathematically seen—the more difficult case as only the real part of the data is available for the inverse problem later on.

Table 2.1: Background optical parameters

State equation	μ_s' mm^{-1}	$\mu_{a,i}$ mm^{-1}	ε $\mathrm{mm}^{-1}\,\mathrm{M}^{-1}$	ρ
Excitation	0.275	0.036	8.35×10^3	0.2
Emission	0.235	0.029	2.81×10^3	0.2

Inside the disc is a circular fluorescent object with a concentration of $1 \times 10^{-2}\,\mu\mathrm{M}$. Although the emission fields in figure 2.2 have visually nearly the same appearance regardless of the source position, the peak amplitude of the emitted field is strongly dependent on the distance of the source to the inclusion. From this dependence, the distribution of the fluorophore can be reconstructed which is shown in the next chapter.

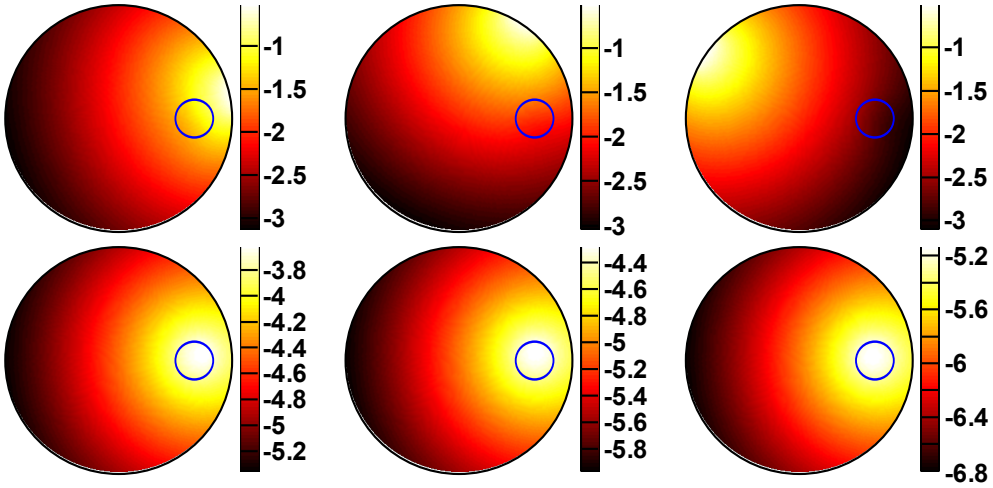

Figure 2.2: Simulation of the photon densities φ_x (top row) and φ_m (bottom row) due to three different boundary sources positioned at angles of 11.25° (left), 56.25° (middle) and 146.25° (right). The colour encodes the logarithm of the amplitude of the photon densities.

3 Inverse problem

> Ich behaupte aber, daß in jeder besonderen Naturlehre nur so viel eigentliche Wissenschaft angetroffen werden könne, als darin Mathematik anzutreffen ist.
>
> *(Immanuel Kant)*

The inverse problem seeks to reconstruct the distribution of a fluorophore from the measurement of light intensities at the emission wavelength on the boundary of the object. There are various approaches to solving this problem and a good overview about reconstruction methods applied for fluorescence tomography can be found in [19].

3.1 Gauß-Newton algorithm

We start from a boundary measurement $\mathcal{M}^\delta$ which is usually perturbed by noise from various origins, e.g. detector noise, discretisation errors, modelling errors and so on. δ is a bound for the deviation of the measured intensity to the true intensity $\mathcal{M} = F(c)$ such that

$$\|\mathcal{M}^\delta - \mathcal{M}\| \leq \delta. \tag{3.1}$$

The norm over the measurement space $\mathbb{C}^{nd\times ns}$ is induced by the scalar product

$$\langle A, B\rangle = \sum_{ij} A_{ij}\overline{B}_{ij} \tag{3.2}$$

Given a set of boundary measurements $\mathcal{M}^\delta$, the inverse problem is to find a suitable fluorophore concentration $c \in C_{ad}$ such that

$$F(c) = \mathcal{M}^\delta. \tag{3.3}$$

This constitutes a nonlinear ill-posed problem [20]. To deal with the nonlinearity, Newton type methods can be used. Starting with some initial guess c_0, a better approximation c_{k+1} for the fluorophore concentration can be found by solving

$$F'(c_k)(c_{k+1} - c_k) = \mathcal{M}^\delta - F(c_k). \tag{3.4}$$

$F'(c)h$ is the derivative of the forward operator F into direction h evaluated for a given concentration c. Except for the special, and highly unlikely, case where the number of measurements is exactly the same as the number of degrees of freedom for the concentration, (i.e. the number of basis function used in the approximation of C_{ad}), F' is not directly invertible and thus the equation is not well-formed.

Due to these shortcomings, a least-squares minimisation with a Tikhonov-type regularisation is considered for a stable inversion. The objective function which is to be minimised takes the form

$$\mathcal{L}_\alpha(c) := \tfrac{1}{2} \left\| F(c) - \mathcal{M}^\delta \right\|^2 + \alpha \mathcal{R}(c). \tag{3.5}$$

In this functional $\mathcal{R}(c)$ is a regularisation functional which stabilises the inversion by penalising unlikely (or undesired) concentration distributions c. The regularisation parameter α is used to balance between the penalisation of the data misfit by the first term in equation (3.5) and the stability of the inversion. In the extreme case $\alpha = 0$ the reconstructed concentration will fit exactly to the measured data which means that $F(c)$ also incorporates the measurement noise which is of course not desired. The other extreme is to choose a very large regularisation parameter with the consequence that the inversion is very stable but $F(c)$ will not at all reflect the measured intensities.

The first-order optimality condition for (3.5) is formally[1] given by

$$\operatorname{Re} \left\langle F'(c)h, F(c) - \mathcal{M}^\delta \right\rangle + \alpha \mathcal{R}'(c)h = 0 \qquad \forall h \in L^2(\Omega). \tag{3.6}$$

[1] Amongst others, it has to be ensured that the updated concentration $c + h$ is part of C_{ad} for a mathematically correct treatment.

Here $\langle \cdot, \cdot \rangle$ is the standard inner product on the space of measurements $\mathbb{C}^{nd \times ns}$ and Re denotes the real part of the following expression. This is again a nonlinear problem in the concentration c, so Newton's method can be used once more which leads to the iteration scheme

$$\begin{aligned} \operatorname{Re} \left\langle F''(c_k)(h, c_{k+1} - c_k), F(c_k) - \mathcal{M}^\delta \right\rangle + \operatorname{Re} \left\langle F'(c_k)h, F'(c_k)(c_{k+1} - c_k) \right\rangle \\ + \alpha_k \mathcal{R}''(c_k)(h, c_{k+1} - c_k) = \operatorname{Re} \left\langle F'(c_k)h, \mathcal{M}^\delta - F(c_k) \right\rangle - \alpha_k \mathcal{R}'(c_k)h, \end{aligned} \tag{3.7}$$

where we have linearized $\mathcal{R}$ around c_k. Note that the regularisation parameter can vary in every iteration which is denoted as α_k. The idea is to start with a large regularisation parameter in order to find a stable approximation to the solution of the inverse problem. In later iterations, when the approximate solution is near the optimum, smaller parameters can be utilized.

When the Hessian $F''(c_k)$ is neglected, we arrive at the Gauß-Newton approximation

$$\begin{aligned} \operatorname{Re} \left\langle F'(c_k)h, F'(c_k)(c_{k+1} - c_k) \right\rangle + \alpha_k \mathcal{R}''(c_k)(h, c_{k+1} - c_k) \\ = \operatorname{Re} \left\langle F'(c_k)h, \mathcal{M}^\delta - F(c_k) \right\rangle - \alpha_k \mathcal{R}'(c_k)h, \end{aligned} \tag{3.8}$$

which is now linear in the update step $\Delta c_k := c_{k+1} - c_k$. It turns out that the Hessian of the regularisation term $\mathcal{R}''$ can be nasty to compute especially in the case of nonlinear methods like total-variation regularisation, for example. Thus, we choose to approximate it by a computationally more efficient version $\widetilde{\mathcal{R}}''$. By introducing F'^*, the adjoint of the derivative of the forward operator, the final iterations are of the form

$$\operatorname{Re} \left(F'^*(c_k)F'(c_k) + \alpha_k \widetilde{\mathcal{R}}''(c_k) \right) \Delta c_k = \operatorname{Re} F'^*(c_k) \left(\mathcal{M}^\delta - F(c_k) \right) - \alpha_k \mathcal{R}'(c_k) \tag{3.9}$$

This is known as the iteratively regularised Gauß-Newton algorithm and is treated in detail elsewhere [26, 27, 28].

3.2 Sensitivity

In the previous section the derivative F' of the forward operator F was introduced. In the following we define this derivative—also known as the *sensitivity* or the *Jacobian*—and discuss an efficient implementation.

First, we are interested in how the change in the fluorophore distribution is reflected in the measurements. To this end, we start from the forward system defined in (2.5)–(2.8). Next, we consider a small perturbation th with $t \in \mathbb{R}$ and $c + th \in C_{ad}$ in the fluorophore concentration c, i.e. the concentration changes to $c + th$. This change will give rise to perturbed optical parameters $\widetilde{\kappa}_i := \kappa_i(c + th)$ and $\widetilde{\mu}_i := \mu_i(c + th)$, $i \in \{x, m\}$, and further also to perturbed photon fields $\widetilde{\varphi}_i$:

$$-\nabla \cdot (\widetilde{\kappa}_x \nabla \widetilde{\varphi}_x) + \widetilde{\mu}_x \widetilde{\varphi}_x = q_{int} \tag{3.10}$$

$$-\nabla \cdot (\widetilde{\kappa}_m \nabla \widetilde{\varphi}_m) + \widetilde{\mu}_m \widetilde{\varphi}_m = \gamma(c + th)\widetilde{\varphi}_x. \tag{3.11}$$

The absorption coefficients are linear in c and the perturbed absorption coefficients are therefore

$$\widetilde{\mu}_i = \mu_{a,i} + \varepsilon_i(c + th) + \frac{i\omega}{\nu}, \qquad i \in \{x, m\}. \tag{3.12}$$

The diffusion coefficients depend on the concentration c in a nonlinear manner. Thus, a Taylor expansion around c of the form

$$\widetilde{\kappa}_i = \kappa_i(c + th) = \kappa_i(c) - \kappa_i^2(c)\, N\varepsilon_i th + \mathcal{O}(t^2\|h\|^2), \qquad i \in \{x, m\}, \tag{3.13}$$

is used. The second order term $\mathcal{O}(t^2\|h\|^2)$ will vanish eventually when taking the limit $t \to 0$ later on which is why this term is neglected right now. The change in the photon fields is the difference between the perturbed and unperturbed fields. This leads to the equations

$$-\nabla \cdot (\kappa_x \nabla(\widetilde{\varphi}_x - \varphi_x)) + \mu_x(\widetilde{\varphi}_x - \varphi_x) = -\nabla \cdot \left(\kappa_x^2\, N\varepsilon_x th \nabla \widetilde{\varphi}_x\right) + (\mu_x - \widetilde{\mu}_x)\widetilde{\varphi}_x, \tag{3.14}$$

$$\begin{aligned} -\nabla \cdot (\kappa_m \nabla(\widetilde{\varphi}_m - \varphi_m)) + \mu_m(\widetilde{\varphi}_m - \varphi_m) = &-\nabla \cdot \left(\kappa_m^2\, N\varepsilon_m th \nabla \widetilde{\varphi}_m\right) + (\mu_m - \widetilde{\mu}_m)\widetilde{\varphi}_m \\ &+ \gamma c(\widetilde{\varphi}_x - \varphi_x) + \gamma th \widetilde{\varphi}_x, \end{aligned} \tag{3.15}$$

and after some reordering and division by t one obtains

$$-\nabla\cdot\left(\kappa_x\nabla\left(\frac{\widetilde{\varphi}_x-\varphi_x}{t}\right)\right)+\mu_x\left(\frac{\widetilde{\varphi}_x-\varphi_x}{t}\right)=-\nabla\cdot\left(\kappa_x^2\,N\varepsilon_x h\nabla\widetilde{\varphi}_x\right)-\varepsilon_x h\widetilde{\varphi}_x, \tag{3.16}$$

$$\begin{aligned}-\nabla\cdot\left(\kappa_m\nabla\left(\frac{\widetilde{\varphi}_m-\varphi_m}{t}\right)\right)+\mu_m\left(\frac{\widetilde{\varphi}_m-\varphi_m}{t}\right)=&-\nabla\cdot\left(\kappa_m^2\,N\varepsilon_m h\nabla\widetilde{\varphi}_m\right)-\varepsilon_m h\widetilde{\varphi}_m\\&+\gamma c\left(\frac{\widetilde{\varphi}_x-\varphi_x}{t}\right)+\gamma h\widetilde{\varphi}_x.\end{aligned} \tag{3.17}$$

The same procedure is applied to the boundary conditions given in equations (2.7)–(2.8) which results in the expression

$$\varrho_i\left(\frac{\widetilde{\varphi}_i-\varphi_i}{t}\right)+\kappa_i n\cdot\nabla\left(\frac{\widetilde{\varphi}_i-\varphi_i}{t}\right)=\kappa_i^2\,N\varepsilon_i hn\cdot\nabla\widetilde{\varphi}_i,\qquad i\in\{x,m\}. \tag{3.18}$$

Finally, we compute the limit $t\to 0$. Here the assumption has to be made that the photon fields depend continuously on c such that $\widetilde{\varphi}_i=\varphi_i(c+th)\overset{t\to0}{\to}\varphi_i(c)$. The infinitesimal changes in the photon fields, induced by a concentration perturbation h, are abbreviated by

$$w_i:=\lim_{t\to0}\frac{\widetilde{\varphi}_i-\varphi_i}{t},\qquad i\in\{x,m\}. \tag{3.19}$$

The result is a new system of coupled partial differential equations which is called the *sensitivity system*. It has the form

$$-\nabla\cdot(\kappa_x\nabla w_x)+\mu_x w_x=-\nabla\cdot\left(\kappa_x^2\,N\varepsilon_x h\nabla\varphi_x\right)-\varepsilon_x h\varphi_x,\qquad\text{in }\Omega, \tag{3.20}$$

$$\begin{aligned}-\nabla\cdot(\kappa_m\nabla w_m)+\mu_m w_m=&-\nabla\cdot\left(\kappa_m^2\,N\varepsilon_m h\nabla\varphi_m\right)-\varepsilon_m h\varphi_m\\&+\gamma c w_x+\gamma h\varphi_x,\qquad\text{in }\Omega,\end{aligned} \tag{3.21}$$

with the boundary conditions

$$\varrho_x w_x+\kappa_x n\cdot\nabla w_x=\kappa_x^2\,N\varepsilon_x hn\cdot\nabla\varphi_x,\qquad\text{on }\partial\Omega, \tag{3.22}$$

$$\varrho_m w_m+\kappa_m n\cdot\nabla w_m=\kappa_m^2\,N\varepsilon_m hn\cdot\nabla\varphi_m,\qquad\text{on }\partial\Omega. \tag{3.23}$$

Based on the definition of the measurement in (2.13), the perturbed measurement

is given by

$$\widetilde{\mathcal{M}}_{ij} = -\int_{\partial\Omega} \left(\kappa_m - \kappa_m^2 \, N\varepsilon_m th\right) n \cdot \nabla\widetilde{\varphi}_m(q_j) d_i \, ds \tag{3.24}$$

and consequently the change in the measurement is

$$\begin{aligned}
\mathcal{M}'_{ij} :&= \lim_{t\to 0} \frac{\widetilde{\mathcal{M}}_{ij} - \mathcal{M}_{ij}}{t} \\
&= -\int_{\partial\Omega} \kappa_m n \cdot \nabla w_m(q_j) d_i \, ds + \int_{\partial\Omega} \kappa_m^2 \, N\varepsilon_m h n \cdot \nabla\varphi_m(q_j) d_i \, ds \\
&= \varrho_m \int_{\partial\Omega} w_m(q_j) d_i \, ds,
\end{aligned} \tag{3.25}$$

which is just the integral of the change in the photon fields over the detector aperture defined by d_i, a result which could also be deduced intuitively.

Given a perturbation h, the sensitivity system (3.20)–(3.23) has to be solved for every source q_j, $j = 1, \ldots, ns$, which enters the system via the photon fields φ_x and φ_m. Then, the induced change in the measurements can be evaluated for every detector d_i, $i = 1, \ldots, nd$, using the representation (3.25). Thus, one arrives at a mapping from the perturbation h around the current distribution c to the variation in the detector recordings which is just the directional derivative $F'(c)h$.

The sensitivity F' maps a given perturbation h to the change in the measurements. In order to solve the inverse problem, the opposite is required, i.e. one can compute the difference between the predicted detector readings $F(c)$ and the true measurement matrix $\mathcal{M}^\delta$ and aims to deduce the perturbation h which caused this change. After discretisation, the number of basis functions for the admissible fluorophore space C_{ad} is finite. Therefore, it would be possible (at least in principle) to solve the system (3.20)–(3.23) for basis function $h \in L^2(\Omega)$. The outcome is a tensor with dimension $nd \times ns \times nc$ or a matrix of size $(nd \cdot ns) \times nc$ where nc is the number of degrees of freedom for C_{ad}. The sought-after perturbation h can be obtained after a matrix inversion. Obviously it is a very time-consuming task to solve one PDE system for every basis function. This motivates the introduction of the adjoint system to devise an efficient strategy for the assembly of the sensitivity matrix.

3.3 Adjoints

The first step is to re-write the sensitivity system (3.20)–(3.23) in a weak form. The equations for w_x and w_m for a certain source q_j are tested with ψ_x and ψ_m, respectively, which gives

$$\begin{aligned}\int_\Omega \kappa_x \nabla w_x \cdot \nabla \overline{\psi}_x dx + \int_\Omega \mu_x w_x \overline{\psi}_x dx + \varrho_x \int_{\partial\Omega} w_x \overline{\psi}_x ds \\ = \int_\Omega \kappa_x^2 N \varepsilon_x h \nabla \varphi_x \cdot \nabla \overline{\psi}_x dx - \int_\Omega \varepsilon_x h \varphi_x \overline{\psi}_x dx \qquad (3.26)\end{aligned}$$

$$\begin{aligned}\int_\Omega \kappa_m \nabla w_m \cdot \nabla \overline{\psi}_m dx + \int_\Omega \mu_m w_m \overline{\psi}_m dx + \varrho_m \int_{\partial\Omega} w_m \overline{\psi}_m ds - \int_\Omega \gamma c w_x \overline{\psi}_m dx \\ = \int_\Omega \kappa_m^2 N \varepsilon_m h \nabla \varphi_m \cdot \nabla \overline{\psi}_m dx - \int_\Omega \varepsilon_m h \varphi_m \overline{\psi}_m dx + \int_\Omega \gamma h \varphi_x \overline{\psi}_m dx \qquad (3.27)\end{aligned}$$

So far, the test functions ψ_x, ψ_m are arbitrary. Now, they will be fixed through the requirement that equation (3.25) and the sum of the left-hand sides of the equations (3.26)–(3.27) should be equal. The reason for this requirement is that it will finally lead to an efficient formulation of the sensitivity system, i.e. the choice of the test functions is motivated by the outcome of the following calculations. The equality requirement stated previously reads

$$\begin{aligned}\int_\Omega \kappa_m \nabla w_m \cdot \nabla \overline{\psi}_m dx + \int_\Omega \mu_m w_m \overline{\psi}_m dx + \varrho_m \int_{\partial\Omega} w_m \overline{\psi}_m ds - \int_\Omega \gamma c w_x \overline{\psi}_m dx, \\ + \int_\Omega \kappa_x \nabla w_x \cdot \nabla \overline{\psi}_x dx + \int_\Omega \mu_x w_x \overline{\psi}_x dx + \varrho_x \int_{\partial\Omega} w_x \overline{\psi}_x ds \stackrel{!}{=} \varrho_m \int_{\partial\Omega} w_m d_i \, ds, \qquad (3.28)\end{aligned}$$

which is after one more partial integration

$$\begin{aligned}- \int_\Omega \nabla \cdot \left(\kappa_m \nabla \overline{\psi}_m\right) w_m dx + \int_\Omega \mu_m w_m \overline{\psi}_m dx - \int_\Omega \gamma c w_x \overline{\psi}_m dx \\ + \varrho_m \int_{\partial\Omega} w_m \overline{\psi}_m ds + \int_{\partial\Omega} \kappa_m n \cdot \nabla \overline{\psi}_m w_m ds - \int_\Omega \nabla \cdot \left(\kappa_x \nabla \overline{\psi}_x\right) w_x dx + \int_\Omega \mu_x w_x \overline{\psi}_x dx \\ + \varrho_x \int_{\partial\Omega} w_x \overline{\psi}_x ds + \int_{\partial\Omega} \kappa_x n \cdot \nabla \overline{\psi}_x w_x ds = \varrho_m \int_{\partial\Omega} w_m d_i \, ds. \qquad (3.29)\end{aligned}$$

This equation has to be fulfilled for every w_x, w_m so it is the weak form of the

system

$$-\nabla \cdot (\kappa_m \nabla \psi_m) + \overline{\mu}_m \psi_m = 0 \qquad \text{in } \Omega \tag{3.30}$$

$$-\nabla \cdot (\kappa_x \nabla \psi_x) + \overline{\mu}_x \psi_x = \overline{\gamma} c \psi_m \qquad \text{in } \Omega \tag{3.31}$$

$$\varrho_m \psi_m + \kappa_m n \cdot \nabla \psi_m = \varrho_m d_i \qquad \text{on } \partial\Omega \tag{3.32}$$

$$\varrho_x \psi_x + \kappa_x n \cdot \nabla \psi_x = 0 \qquad \text{on } \partial\Omega, \tag{3.33}$$

which is the *adjoint system* to the forward system (2.5)–(2.8). We note that this adjoint system has to be solved in reverse order: The detector is the source for the adjoint emission photons ψ_m which propagate through the tissue. Where a fluorophore is present, the absorbed emission photons will be converted in part to photons at the excitation wavelength, which in turn spread through the tissue again. Thus, one can say that the adjoint system "reverses" the physical process of light propagation. We write $\psi_x(d_i)$ and $\psi_m(d_i)$ to emphasize the dependence of the adjoint fields on the detector d_i.

Finally, we can replace the left-hand side of (3.28) by the right-hand sides of (3.26)–(3.27). Then the sensitivity for the ij-th measurement is given by

$$\begin{aligned}
\mathcal{M}'_{ij} =& \varrho_m \int_{\partial\Omega} w_m(q_j) d_i \, ds \\
=& \int_\Omega \kappa_x^2 \, N \varepsilon_x h \nabla \varphi_x(q_j) \cdot \nabla \overline{\psi}_x(d_i) dx - \int_\Omega \varepsilon_x h \varphi_x(q_j) \overline{\psi}_x(d_i) dx \\
&+ \int_\Omega \kappa_m^2 \, N \varepsilon_m h \nabla \varphi_m(q_j) \cdot \nabla \overline{\psi}_m(d_i) dx - \int_\Omega \varepsilon_m h \varphi_m(q_j) \overline{\psi}_m(d_i) dx \\
&+ \int_\Omega \gamma h \varphi_x(q_j) \overline{\psi}_m(d_i) dx.
\end{aligned} \tag{3.34}$$

Instead of the cumbersome assembly of the sensitivity matrix described in the direct approach at the end of 3.2, we have a much more powerful method now. It suffices to solve the forward problem (3.20)–(3.23) once for every source q_j, $j = 1, \ldots, ns$, and the adjoint problem (3.30)–(3.33) once for every detector d_i, $i = 1, \ldots, nd$. The sensitivity matrix can then be assembled using the representation (3.34). With this approach only $(ns + nd)$ systems of partial differential equations have to be solved instead of solving one PDE system per basis function as in the direct approach.

The only quantity left for explanation in equation (3.9) is the adjoint sensitivity F'^*.

It is defined through the equality of the inner products

$$\langle F'(c)h, r\rangle \stackrel{!}{=} \langle h, F'^{*}(c)r\rangle\,. \tag{3.35}$$

The inner product on the left-hand side is over the measurement space $\mathbb{C}^{nd\times ns}$ while the one on the right-hand side is the one induced by the L^2 function space. Writing down the inner product on the left-hand side and using (3.34) results in

$$\begin{aligned}\sum_{ij} \varrho_m \int_{\partial\Omega} w_m(q_j) d_i \, ds \, \overline{r}_{ij} \\ = \sum_{ij} \int_{\Omega} h \Big[\kappa_x^2 \, N \varepsilon_x \nabla\varphi_x(q_j)\cdot\nabla\overline{\psi}_x(d_i) - \varepsilon_x\varphi_x(q_j)\overline{\psi}_x(d_i) \\ + \kappa_m^2 \, N \varepsilon_m \nabla\varphi_m(q_j)\cdot\nabla\overline{\psi}_m(d_i) - \varepsilon_m\varphi_m(q_j)\overline{\psi}_m(d_i) \\ + \gamma\varphi_x(q_j)\overline{\psi}_m(d_i)\Big]\overline{r}_{ij}\, dx,\end{aligned} \tag{3.36}$$

following that the functional $F'^{*}(c)$ can be identified with the complex conjugate of the expression in the square brackets.

3.4 Discretisation

The adjoint system is discretised in the same manner as the forward model in 2.3. The finite element system has the form

$$\left[K(\kappa_m) + M(\overline{\mu}_m) + B(\varrho_m)\right]\psi_m = d, \tag{3.37}$$

$$\left[K(\kappa_x) + M(\overline{\mu}_x) + B(\varrho_x)\right]\psi_x = M(\overline{\gamma}c)\psi_m, \tag{3.38}$$

i.e. the same structure is obtained as for the forward system with the notable exceptions that (i) the order of the excitation and emission matrix is reversed and (ii) the complex coefficients μ_x, μ_m and γ have to be conjugated.

The sensitivity matrix $F'(c)$ is of size $(nd\cdot ns)\times nc$. The action of the ij-th row,

$1 \leq i \leq nd$, $1 \leq j \leq ns$, on a parameter perturbation $h \in \mathbb{R}^{nc}$ can be written as

$$\begin{aligned} F'_{ij}h := &-\overline{\psi}_m(d_i)^\top K(\kappa'_m h)\varphi_m(q_j) - \overline{\psi}_m(d_i)^\top M(\mu'_m h)\varphi_m(q_j) \\ &- \overline{\psi}_x(d_i)^\top K(\kappa'_x h)\varphi_x(q_j) - \overline{\psi}_x(d_i)^\top M(\mu'_x h)\varphi_x(q_j) + \overline{\psi}_m(d_i)^\top M(\gamma h)\varphi_x(q_j), \end{aligned} \tag{3.39}$$

where κ'_x, κ'_m, μ'_x and μ'_m are the derivatives of the coefficients of the governing equations with respect to the concentration c. For an efficient evaluation, the diffusion coefficients are projected onto the finite element space of piecewise constant functions prior to the computation of the derivative. This is not needed for the absorption parameters as they are linear in the concentration. For a finite element T, the derivatives are given by the expressions

$$\kappa'_{i,T} = -N\epsilon_i \left(\frac{1}{4} \sum_{j=1}^{4} \kappa_{i,g_T(j)} \right)^2 \qquad \text{and} \qquad \mu'_i = \epsilon_i, \qquad i \in \{x, m\}. \tag{3.40}$$

The discretized adjoint derivative $F'^*(c)$ is another matrix of size $nc \times (nd \cdot ns)$. Using (3.35) it is realised as

$$F'^*(c) = M(1)^{-1}\overline{F'}^\top(c), \tag{3.41}$$

where the mass matrix $M(1)$ incorporates the change of the underlying spaces used for the inner products.

3.5 Examples

An example of the appearance of the adjoint fields for the same circular perturbation as in figure 2.2 and three different detector positions is shown in figure 3.1. As the PDE system is similar to the forward system, the appearance of the result is also similar.

The sensitivity pattern for chosen source and detector locations is shown in figure 3.2. One often refers to a *banana-shape* when characterising the appearance of these "sensitivity fields". A prototype example of this shape can be seen in the second column of the last row.

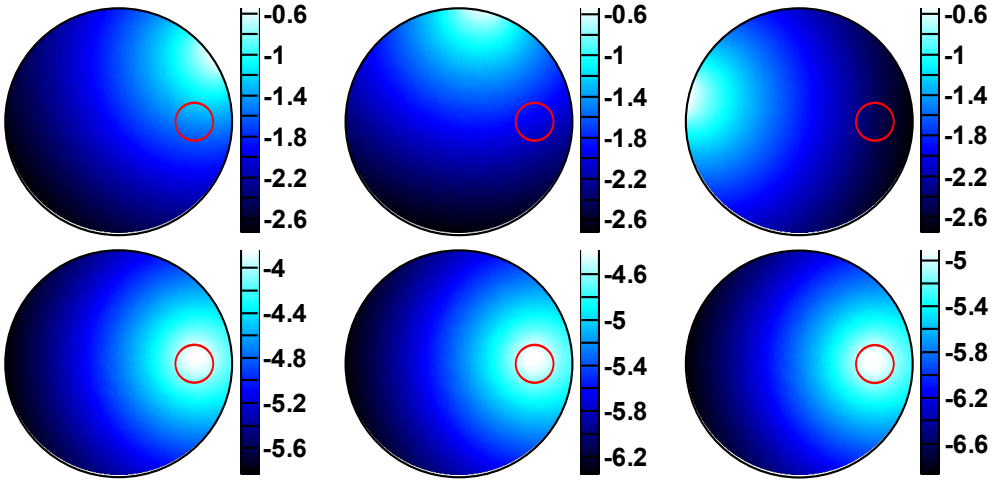

Figure 3.1: Distribution of the logarithm of the adjoint fields ψ_m (top row) and ψ_x (bottom row) due to three different detector positions at angles of 33.75° (left), 78.75° (middle) and 168.75° (right).

One notices that the magnitude of the sensitivity fields drops with increasing sensor/detector separation. Therefore, the signal of adjacent source/detector pairs is at least an order of a magnitude larger than those stemming from transillumination measurements with opposing optodes.

The measurement of an optode pair is the volume integral of the fluorophore concentration weighted by the corresponding sensitivity. Because the sensitivity is quite smooth due to the diffusive nature of photon propagation in highly scattering media, every measurement will contain information from a considerable fraction of the total volume. Thus, it is immediately clear that the resolution of fluorescence tomography is rather poor. This is also the source of the ill-posedness of the inverse problem which makes the use of regularisation methods mandatory. Possible choices for the regularisation terms are treated in the next chapter.

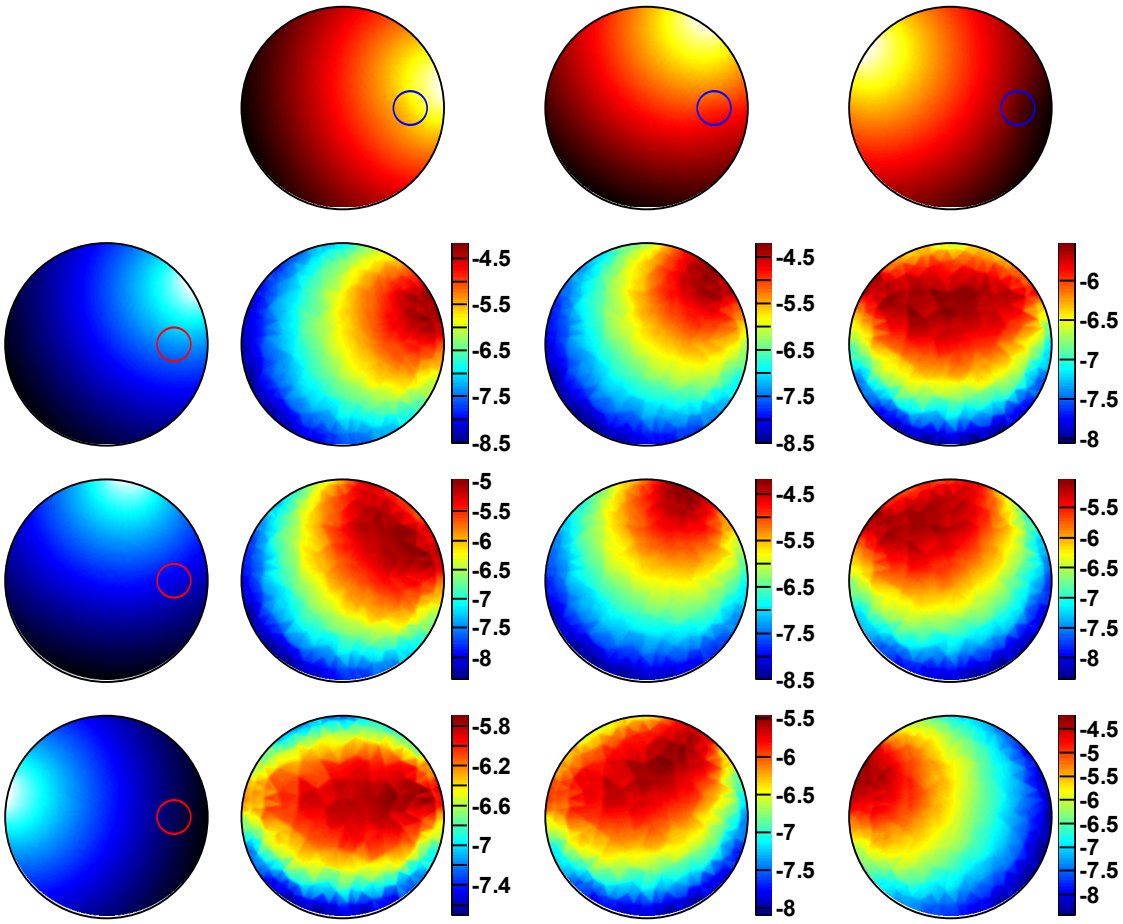

Figure 3.2: Logarithm of the rows of the sensitivity matrix for different source and detector locations. For every sensitivity field the corresponding source is displayed in the first row and the associated detector in the column farthest left.

4 Problem-adapted regularisation

> Die Rechenautomaten haben etwas von den Zauberern im Märchen. Sie geben einem wohl, was man sich wünscht, doch sagen sie einem nicht, was man sich wünschen soll.
>
> *(Norbert Wiener)*

In a typical fluorescence tomography setup one has more degrees of freedom, i.e. more voxels in the computation domain, than measurements. So the reconstruction of the fluorophore concentration c is underdetermined. Therefore, one can tune the concentration in the voxels in such a way that the data misfit $\|F(c) - \mathcal{M}\|$ is zero. Then the concentration c would not only reflect the intensity measurements but also the noise in them. In fact, there are even infinitely many possibilities for a concentration distribution with a vanishing data misfit. See the book by Tarantola [29] for a more rigorous discussion of this problem.

Obviously, one needs to select a suitable fluorophore distribution c from all possible ones. In order to do so, one has to bring a-priori information into the reconstruction algorithm. One possibility to do so is through the choice of the regularisation functional $\mathcal{R}$ in equation (3.5).

The following chapter is devoted to the introduction of quadratic and nonquadratic regularisation terms and a comparison of the resultant reconstructions.

4.1 Quadratic regularisation terms

The most often used regularisation term is the L^2 norm of the concentration. This term penalises large concentrations and, thus, favours lower concentrations in the whole domain. The dominance of such a penalty in many applications is not due to its physical meaning but rather due to its simplicity and speed, both in implementation and evaluation. The regularisation term and its derivatives are

$$\mathcal{R}(c) = \tfrac{1}{2}\|c\|^2, \qquad \mathcal{R}'(c) = c, \qquad \mathcal{R}''(c) = I, \tag{4.1}$$

where I is the identity operator.

When the initial concentration is not $c \equiv 0$, i.e. it is not uniformly zero everywhere, but one can a-priori provide a better guess c_0, it is advantageous to penalise the distance to this suggestion through the usage of the regularisation term

$$\mathcal{R}(c) = \tfrac{1}{2}\|c - c_0\|^2. \tag{4.2}$$

Another possibility is to penalise the gradient of the reconstructed fluorophore distribution. In contrast to an L^2 regularisation, the regularisation by the H^1 semi-norm allows for larger fluorophore concentrations but it penalises fluctuations in the reconstruction and will thus force the resultant image to be smooth. The regularisation functional is defined as

$$\mathcal{R}(c) = \tfrac{1}{2}\|\nabla c\|^2, \qquad \mathcal{R}'(c)h = \langle \nabla c, \nabla h \rangle, \qquad \langle \mathcal{R}''(c)h_1, h_2 \rangle = \langle \nabla h_1, \nabla h_2 \rangle \tag{4.3}$$

4.2 Total-variation regularisation

A disadvantage of the regularisation term defined in (4.3) is that it keeps the gradient of the fluorophore concentration small. Thus, jumps are considered to be "expensive" and are avoided leading to a smooth image instead. For certain applications where the fluorophore concentration is a-priori known to be piecewise constant (e.g. it might be constant in a tumour or an organ and zero in the surrounding), the H^1 reconstruction will be unnecessarily diffusive.

This is mitigated to some extent by the use of a total-variation (TV) regularisation term. Although it also measures the gradient of the reconstructed image, it uses the L^1 norm rather than the L^2 norm as in equation (4.3). The total-variation functional is defined as

$$\mathrm{TV}(c) := \int_\Omega |\nabla c| \, dx. \tag{4.4}$$

This penalty allows jumps in the image while it suppresses oscillations (or variations) in the fluorophore concentration.

The drawback of the TV functional is its non-differentiability at points where $\nabla c = 0$, i.e. where the concentration is constant. Thus, we use a relaxation of the functional of the form [30]

$$\begin{aligned} \mathcal{R}(c) &= \int_\Omega \sqrt{\beta + |\nabla c|^2} dx, \qquad \mathcal{R}'(c)h = \int_\Omega \frac{\nabla c \cdot \nabla h}{\sqrt{\beta + |\nabla c|^2}} \, dx, \\ \left\langle \widetilde{\mathcal{R}}''(c)h_1, h_2 \right\rangle &= \int_\Omega \frac{\nabla h_1 \cdot \nabla h_2}{\sqrt{\beta + |\nabla c|^2}} \, dx \end{aligned} \tag{4.5}$$

with some scalar $\beta > 0$ which smooths the total-variation functional around the origin. Note also that $\widetilde{\mathcal{R}}''$ is just an approximation of the true Hessian, because the derivative of the denominator has been neglected.

4.3 Method of levelset type

Also levelset methods have attracted some interest in the field of diffuse optical tomography for the reconstruction of piecewise constant parameters, see e.g. [31, 32]. The idea is the following: if the background fluorescence and the fluorescence inside the inclusions are constant and known a-priori, the reconstruction will reduce to the determination of the interface which separates the fluorophore inclusions from the surrounding. One method to describe the interface mathematically is through the levelset function ϕ which is defined on the whole domain Ω. All points x with $\phi(x) < 0$ are considered to be inside the inclusion while for locations x in the background it is true that $\phi(x) > 0$. Thus, the set $\{x : \phi(x) = 0\}$, which is the zero levelset, is the definition of the separating boundary (see figure 4.1 for a sketch of this concept).

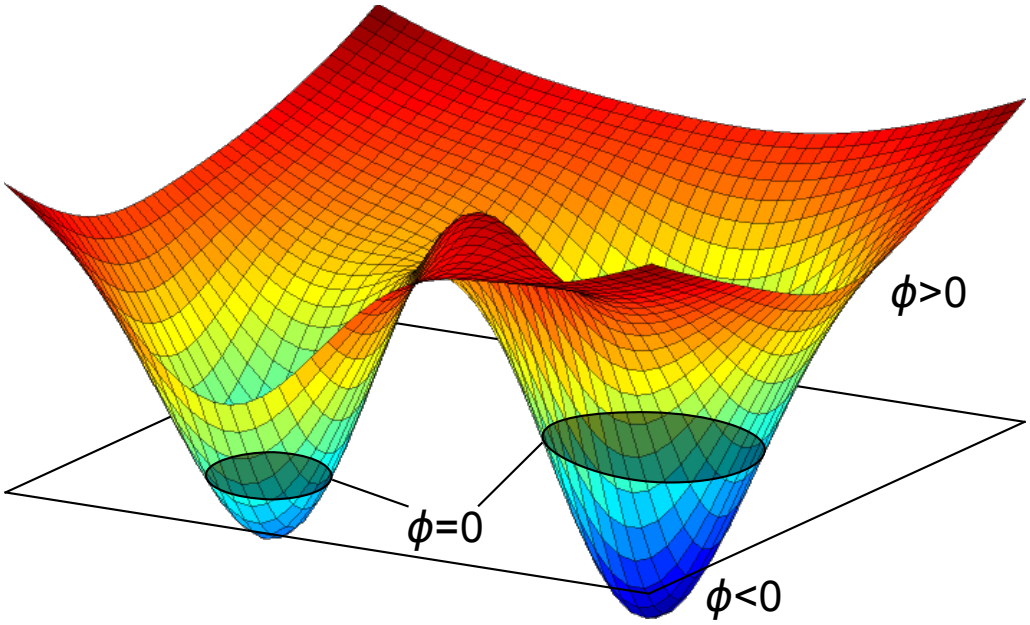

Figure 4.1: A levelset function ϕ which specifies two elliptical inclusions (shaded area) in a rectangular domain Ω. The interior is defined by $\phi(x) < 0$.

Having computed the levelset function, the concentration in any point $x \in \Omega$ can be written as

$$c(x) = c_u + H(\phi(x))(c_l - c_u), \tag{4.6}$$

where c_u and c_l are the upper (inside the inclusions) and lower (background) concentration levels, respectively, and H is the Heaviside function

$$H(x) = \begin{cases} 1, & x \geq 0, \\ 0, & \text{else.} \end{cases} \tag{4.7}$$

Unfortunately, the numerical effort for the computation of the evolution of the levelset function is comparatively high. The reason is again the non-differentiability of the Heaviside function. Thus, we utilise what has been studied as a *method of levelset type* in [33] and [34]. The idea is to parametrise the concentration c by a nonlinear, smooth, monotonically increasing function $H_\beta(\phi)$ and to reconstruct the levelset function ϕ afterwards which is done by minimising the objective function

$$\mathcal{L}_\alpha^\phi = \tfrac{1}{2}\|F(H_\beta(\phi))\|^2 + \alpha\mathcal{R}(\phi). \tag{4.8}$$

In this thesis, we use a smooth parametrisation based on the error function erf. Its definition reads

$$H_\beta(x) := \frac{c_u + c_l}{2} + \frac{c_u - c_l}{2} \operatorname{erf}\left(\frac{x}{\beta}\right). \tag{4.9}$$

This parametrisation restricts the concentration by the lower and upper bounds c_l and c_u, respectively, but still allows values in between. Furthermore, it has a smooth derivative which is needed for the Gauß-Newton algorithm.

The regularisation functional $\mathcal{R}$ now acts on the levelset function ϕ, and we choose the H^1 semi-norm for it.

4.4 Discretisation

The regularisation terms after a finite element discretisation are in the case of L^2-regularisation given by

$$\mathcal{R}'(c) = M(1)c, \qquad \widetilde{\mathcal{R}}''(c) = M(1), \tag{4.10}$$

for H^1-regularisation by

$$\mathcal{R}'(c) = K(1)c, \qquad \widetilde{\mathcal{R}}''(c) = K(1), \tag{4.11}$$

and for the total-variation formulation they read

$$\mathcal{R}'(c) = K(c_\beta)c, \quad \widetilde{\mathcal{R}}''(c) = K(c_\beta), \quad c_\beta := \frac{1}{\sqrt{\beta + |\nabla c|^2}}. \tag{4.12}$$

For the method of levelset type, the regularisation is applied to the levelset function ϕ and has the same appearance as for the concentration above.

4.5 Regularisation parameter and stopping criterion

The choice of the regularisation parameter α in equation (3.5) has attracted quite some interest. The basic problem is that this parameter balances the *goodness of*

fit, i.e. how well the mathematical model approximates the measured data, and the stability of the inversion. In this thesis a geometric sequence $\alpha_k = q^k \alpha_0$ is used for the regularisation parameter. Its initial value α_0 and the factor q are chosen depending on the regularisation method described in sections 4.1, 4.2 and 4.3. An advantage of the iteratively regularised Gauß-Newton algorithm is its robustness to these two parameters [27, 28].

Another problem is the stopping criterion for the iterative reconstruction. Most often these criteria are implemented using heuristics. For example, one can set a lower limit for the regularisation parameter [35] or stop if stagnation occurs and the residual no longer changes [36]. Note that for the latter approach a good choice for the step length is mandatory. Otherwise, the residual will decrease very slowly which triggers the stopping criterion even when no (local) minimum of the objective functional has been reached.

An objective stopping criterion is given by Morozov's discrepancy principle [37]. It states that the residual shall not be smaller than the error in the measurement. Thus, α should be chosen such that

$$\|F(c_\alpha) - \mathcal{M}^\delta\| \approx \|\mathcal{M} - \mathcal{M}^\delta\|, \tag{4.13}$$

where c_α denotes the reconstructed fluorophore distribution with regularisation parameter α. Unfortunately, the expression on the right-hand side cannot be evaluated in practise as the true measurements $\mathcal{M}$ are not known. Thus, one has to deduce this quantity by other means based on the distribution of the noise, for example.

In order to relate Morozov's discrepancy principle to the measurement noise, it is useful to treat the noisy measurements $\mathcal{M}_{ij}^\delta$ as random variables. The measurement consists of two parts: i) the intensity recorded by the detector which is given by the forward operator $F_{ij}(c)$ and ii) some additional noise N_{ij}. We assume our model $F(c)$ to be correct—which means free of systematic errors—and deterministic. The noise N_{ij} consists not only of the detector noise but also of additional errors which are due to simplifications in the forward model, uncertainties of the geometry, the discretisation by finite elements and so on. Then the erroneous measurement has the form

$$\mathcal{M}_{ij}^\delta = F_{ij}(c) + N_{ij}. \tag{4.14}$$

For the sake of simplicity, some constraints are imposed on the measurement noise. Firstly, the noise shall be independent, i.e. the knowledge of one measurement shall not influence the noise of another measurement. Secondly, the noise is assumed to be normally distributed with zero mean and some variance σ_{ij}^2, i.e. $N_{ij} \sim \mathcal{N}(0, \sigma_{ij}^2)$. In case the modelling and discretisation errors are small and the noise is mainly caused by the detectors, it would be physically more sensible to assume a Poisson distribution for N_{ij} as it is the consequence of a photon counting process. However, even in the latter case the normal distribution can be seen as an approximation to the Poisson distribution for larger intensities [38].

If the fluorophore concentration c could be reconstructed perfectly, the elements of the residual matrix R given by

$$R_{ij} := \mathcal{M}_{ij}^{\delta} - F_{ij}(c) = N_{ij}, \tag{4.15}$$

would contain only the noise, and thus $R_{ij} \sim \mathcal{N}(0, \sigma_{ij}^2)$. From this it immediately follows that the scaled residual $\dfrac{R_{ij}}{\sigma_{ij}}$ is distributed as $\mathcal{N}(0, 1)$. Therefore, we stop the iterative Gauß-Newton algorithm defined in (3.9) as soon as the estimated variance of the scaled residual is below 1, i.e. when

$$\operatorname{var}\left(\frac{R}{\sigma}\right) = \frac{1}{nd \cdot ns - 1} \sum_{ij} \left(\frac{R_{ij}}{\sigma_{ij}} - \overline{R}\right)^2 \leq 1, \qquad \text{with } \overline{R} = \frac{1}{nd \cdot ns} \sum_{ij} \frac{R_{ij}}{\sigma_{ij}}. \tag{4.16}$$

The drawback of this simple approach is that it uses the variance of the residuals only and neglects their mean value. So, in certain applications this approach might not work as expected. In these cases it is advantageous to define another random variable Y which is the variance-weighted sum of squares of the residuals

$$Y := \sum_{i=1}^{nd} \sum_{j=1}^{ns} \left(\frac{R_{ij}}{\sigma_{ij}}\right)^2. \tag{4.17}$$

This random variable exhibits a (central) chi-square distribution, $Y \sim \chi^2_{nd \times ns}$, where the number of degrees of freedom is equal to the number of measurements [39]. The mean of this random variable is equal to the degrees of freedom and the variance is twice this value. Therefore, as an alternative to (4.16), one can also stop the iterative

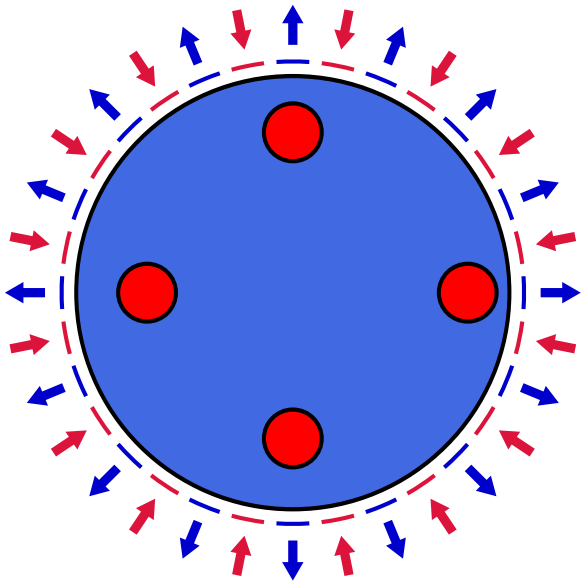

Figure 4.2: Numerical phantom used to test various regularisation methods. The arcs depict the size and location of the optodes where sources are symbolised by inward-pointing and detectors by outward-pointing arrows. The radius of every source and detector was set to 1 mm.

reconstruction as soon as the sum of squares of the weighted residual approaches a chosen threshold t:

$$\sum_{ij} \left(\frac{R_{ij}}{\sigma_{ij}} \right)^2 = \sum_{ij} \left(\frac{\mathcal{M}_{ij}^{\delta} - F_{ij}(c)}{\sigma_{ij}} \right)^2 \leq t. \tag{4.18}$$

4.6 Model problem

The regularisation methods are tested on a two dimensional numerical phantom. It consists of a circular disc with a diameter of 30 mm. 16 sources and the same number of detectors are arranged around the perimeter. Four fluorescent inclusions with a diameter of 4 mm and a fluorophore concentration of 10 μM are placed inside the phantom at different depths. The optical properties of the surrounding are the same as in table 2.1. In counter-clockwise direction starting with the one farthest right, the inclusions' centres are 3 mm, 4 mm, 5 mm and 6 mm beneath the boundary. This setup is shown in figure 4.2.

The finite element mesh consists of 12 289 nodes and 24 064 elements. To avoid inverse crimes [40], the reconstruction is computed on a smaller mesh with only 3137 nodes and 6016 elements. Furthermore, the measurement data is perturbed with Gaussian noise with a standard deviation equal to 1 % of the largest measurement magnitude $\max |\mathcal{M}_{ij}|$.

4.7 Results

The choice of α_0 and q in the geometric sequence $\alpha_k = q^k \alpha_0$ depends on the reconstruction algorithm and is listed in table 4.1. Regarding the method of levelset type, the lower concentration is always set to $c_l = 0\,\mu\text{M}$ while different values are used for the upper level c_u. The stopping index is chosen according to (4.16) or (4.18).

Table 4.1: Regularisation parameters used for testing the reconstruction algorithms.

Parameter	L^2	H^1	TV	erf $+H^1$
α_0	10^{-4}	10^{-4}	10^{-4}	10^{-3}
q	0.5	0.5	0.5	0.5
β	–	–	10^{-16}	0.5

The reconstruction results using the stopping criterion devised in (4.16) are shown in figure 4.3. The estimated concentration inside the inclusions is nearly the same for the L^2-, H^1- and TV-regularised images 4.3(b)–(d). All exhibit increased inclusion areas with a reduced peak amplitude that is only 60 % of the original value. In the L^2 reconstruction artifacts are visible, especially near the boundary of the domain. Such artifacts are reduced in the H^1- and even more in the TV-regularised images. Comparing figures (c) and (b), the smoothing behaviour of the H^1 seminorm due to the penalisation of the gradient can be seen very well. This smoothing is not present in the TV-regularised (figure (d)) reconstruction where the transition from an inclusion to the background is much steeper. Furthermore, one notices also that the total-variation regularisation promotes a constant concentration inside the inclusions. The levelset reconstructions in 4.3(e)–(f) have the best localisation of the perturbations. This comes at the cost of additional a-priori information which is the knowledge of the upper fluorophore concentration c_u. If this quantity can be

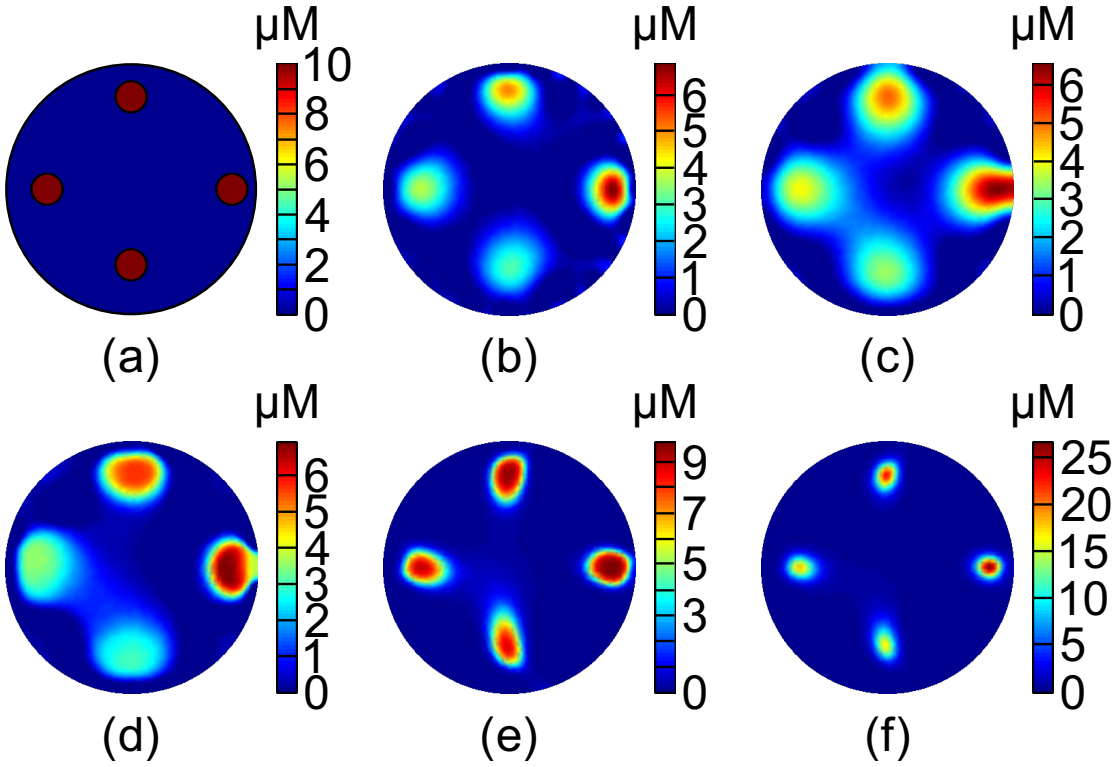

Figure 4.3: Comparison of reconstructions of (a) four fluorescent objects using (b) L^2, (c) H^1 and (d) total-variation regularisation and a method of levelset type with (e) $c_u = 10\,\mu$M and (f) $c_u = 50\,\mu$M. All concentrations are projected to $\mathbb{R}^{nc}_+$. The stopping index is chosen according to the variance approach defined in (4.16).

estimated correctly, the position and size of the inclusions can be reconstructed significantly better than with the other regularisation methods. However, this method is sensitive to the upper concentration as can be seen from figure (f). There, the fluorophore concentration in the perturbations is overestimated by a factor of 2.5 which in turn is the reason for the underestimated area of the inclusions.

For comparison reasons we also show the result of the reconstructed fluorophore when the stopping index is chosen based on the variance-weighted sum of squares given in (4.18). The threshold was set to twice the mean value of the underlying chi-square distribution. The results shown in figure 4.4 are quite similar to the ones obtained with the variance approach shown in figure 4.3. However, the reconstruction is stopped in an earlier iteration when the sum of squares-method is used to choose the stopping index. This is why the method of levelset type (4.4(f)) does not overestimate the fluorophore concentration as much as in 4.3(f).

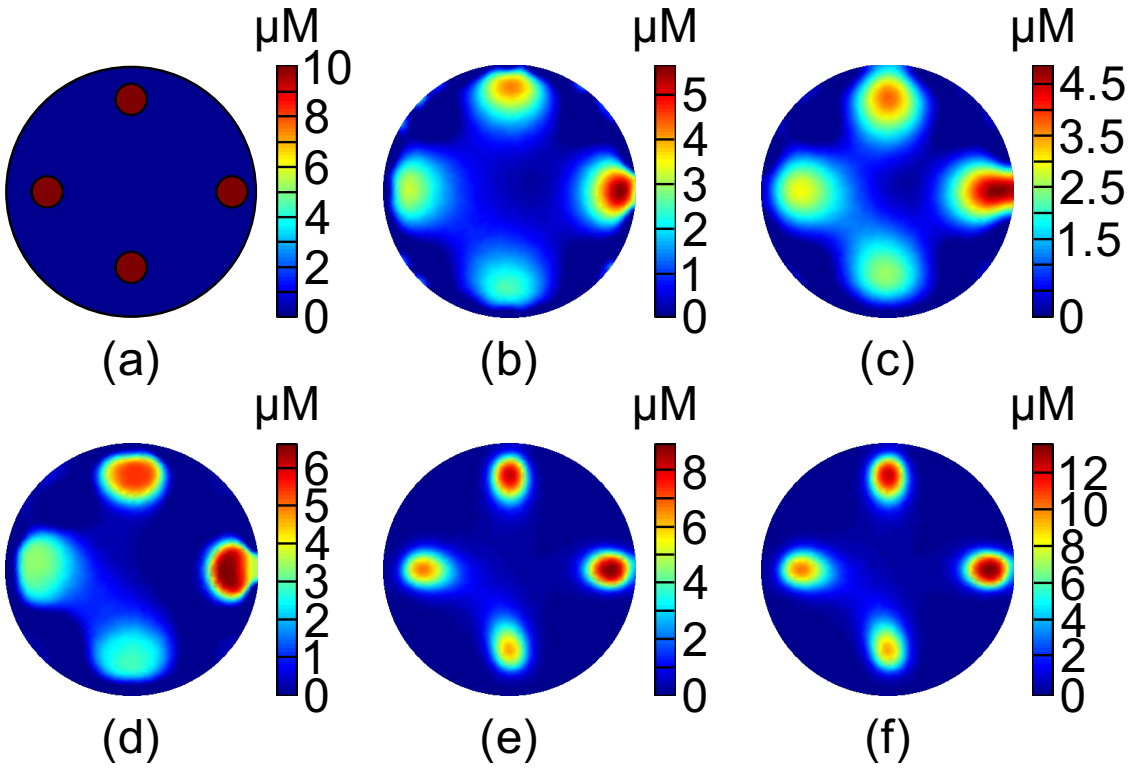

Figure 4.4: Reconstructions from simulated data of the original phantom (a) with (b) L^2, (c) H^1 and (d) TV regularisation and a levelset type method with upper concentration levels of (e) $c_u = 10\,\mu\text{M}$ and (f) $c_u = 50\,\mu\text{M}$. The concentrations are projected to $\mathbb{R}_+$. The variance-weighted sum of squares of (4.18) is used to determine the stopping index.

Further investigations in the reconstruction quality of the previously described regularisation methods have been subject to a study published in [35]. In a cylindrical simulation phantom with a diameter of 30 mm and a height of 60 mm a random number of inclusions with a random Gaussian shaped fluorophore distribution have been placed. In order to compare the result in an objective manner, the amount of reconstructed fluorophore within regions containing $x\,\%$ of the real inclusion is considered. While smaller values for the percentage x provide information about the localisation of the reconstructed fluorophore inclusion, larger percentages are helpful to compare the total amount of fluorophore.

The results of an extensive simulation study with 100 different fluorophore distributions are listed in table 4.2. The heading *linearised* L^2 refers to a single-step reconstruction similar to the Born approximation. This linearisation performs worst for all regions. Similar to the results shown in figures 4.3 and 4.4, total-variation regularisation is more accurate than the ordinary L^2 regularisation. Also in this study, the method of levelset type tended to overestimate the true amount of fluorophore while it is still the method with the best localisation (i.e. most of the fluorophore is located within the 50 % region of interest).

Table 4.2: Ratio of fluorophore in the reconstructions to the original amount for different regularisation strategies.

Region	linearized L^2	L^2	TV	erf $+H^1$
50 %	0.51± 0.12	0.65±0.17	0.80±0.21	1.03±0.50
75 %	0.60± 0.11	0.75±0.16	0.89±0.18	1.18±0.64
90 %	0.69± 0.11	0.86±0.14	0.96±0.15	1.35±0.83
95 %	0.75± 0.11	0.91±0.13	1.00±0.13	1.45±0.97

4.8 Discussion

In this chapter, nonlinear reconstruction schemes for fluorescence tomography with quadratic and nonquadratic regularisation terms have been compared. All regularisation methods be it L^2, H^1, total-variation or the method of levelset type lead to very similar algorithms after discretisation and can therefore be easily incorporated into existing finite element codes.

The iteratively regularised Gauß-Newton reconstruction method uses a decaying sequence of regularisation parameters α_k and does not require potentially expensive line searches which are typically required to ensure global convergence of Newton-type methods. Furthermore, the iterative reconstruction easily allows termination of the reconstruction when a certain a-posteriori criterion is fulfilled. We based the stopping criterion on Morozov's discrepancy principle which has an intuitive physical interpretation—the noise in the reconstruction should not be smaller than the noise in the measurement. However, other criteria that measure the change of the residual, for example, are also possible and in use, cf. [36].

The method of levelset type needs the lower and upper fluorophore concentration levels c_l and c_u. While the lower level for the background can arguably be fixed to $c_l = 0\,\mathrm{M}$, the choice of the upper level is a considerable problem for practical applications. In [35] the approach was tested to relate the peak concentration to the total amount of fluorophore injected into the sample by assuming an average volume for the inclusion. Unfortunately, this assumption does not hold in many cases, which is the reason why the method of levelset type overestimates the amount of fluorophore, as is visible in table 4.2. A possible solution could be to include the determination of the upper concentration level in the reconstructions, i.e. after finding a levelset function, which segments the domain into inclusions and background, another minimisation with respect to the fluorophore concentration inside the inclusion could be performed. These two optimisation steps could be applied in an alternating manner.

It can be concluded that the use of nonquadratic regularisation terms improves the quality of the reconstructions significantly, both visually and also in an objective manner. The method of levelset type is the best choice if information on the expected concentration levels is available a-priori or can be estimated during the reconstruction process. In other cases, the total-variation regularisation might be a viable alternative.

5 Noise considerations

> The Guide is definitive. Reality is frequently inaccurate.
>
> *(Douglas Adams)*

In the previous chapters the inverse problem and nonquadratic regularisation methods have been presented and a method for the choice of the stopping criterion based on the measurement noise has been proposed. In the following the origin of the measurement noise and its influence on the reconstruction will be investigated.

5.1 Origin of noise

The charge-coupled device (CCD) chip used in many scientific cameras is a complex high-tech product. Accordingly, there are many different noise sources and the analysis of CCD noise has been subject to various investigations, e.g. [41, 42, 43].

A short characteristic of important noise types following [41] is given here. Traversing the chain from the arrival of photons on the CCD to the final digital signal, the main noise sources are shot noise, dark noise, amplifier noise and read-out noise.

Shot noise: Shot noise or photon noise is due to the inherent quantum nature of light and therefore cannot be eliminated. The interval between the arrival of photons is governed by a Poisson statistics from which follows that the noise variance is equal to the mean value of the signal.

Dark noise: Dark noise originates in the thermal generation of electrons in the silicon of the integrated circuit. Similar to shot noise, also the dark noise follows a Poisson statistics. However, as it depends heavily on the temperature of the CCD, cooling the camera significantly reduces the amount of thermal electrons.

Amplifier noise: The charge generated by the photons in the bins of the CCD is amplified and converted in a voltage prior to any further signal processing. Thus, the signal is perturbed further by amplifier noise which can be assumed to be normally distributed with zero mean. For low signal intensities, amplifier noise dominates the shot noise. It is important to note that the amplification of the charge generated by the photons also amplifies shot and dark noise. Therefore, cooling the CCD to reduce dark noise is important for low-signal applications.

Read-out noise: The conversion of the analog voltage signal to a digital signal at the read-out stage introduces read-out noise or quantisation noise. This error depends on the number of quantisation steps and the quality of the analog-to-digital converter. This noise can be assumed to be uniformly distributed in $\pm\frac{1}{2}q$, where q is the quantisation step.

5.2 Effect on reconstruction quality

For the reconstruction shown in figures 4.3 and 4.4 the same noise standard deviation of 1 % of the largest measurement was assumed. In the following, the influence of the noise on the reconstructions will be studied in more detail.

The final probability density function for the overall measurement noise is quite complicated as it contains Poisson distributed photon and dark noise, Gaussian amplifier noise and uniform read-out noise. For the sake of simplicity, in this thesis all the noise added to the simulation data is assumed to be normally distributed. This can be justified as for low-intensity signals the detection hardware is the main noise source. For high-intensity signals the Poisson distribution can be approximated by a normal distribution sufficiently well [38].

We test the regularisation methods with noise that is subject to three different assumptions: (N1) a constant standard deviation for all measurements, $\sigma_{ij} = \sigma = f \cdot \max_{ij} |\mathcal{M}_{ij}|$, (N2) a constant standard deviation for all measurements belonging to the same source, $\sigma_{ij} = \sigma_j = f \cdot \max_i |\mathcal{M}_{ij}|$ and (N3) a standard deviation which is proportional to the corresponding detector reading $\sigma_{ij} = f \cdot |\mathcal{M}_{ij}|$. f denotes a scalar fraction and is in the range of several percent.

The first approach is often found in literature and is a sort of a worst case because small signals will suffer considerably. To relate this noise model to a practical setup, several assumptions have to be made: (i) The measurement noise should depend mainly on the integration time of the detector. (ii) The integration time in turn is chosen such that the largest measurement fits in the detector's dynamic range. (iii) The same integration time is used for all measurements.

Next we again assume that the integration time determines the measurement noise. However, we allow this time to be adjusted for every excitation separately rather than having one fixed integration time for all sources. So we end up with noise model (N2). There the variance is constant for all measurements belonging to the same source. To realise such a setup in practice, it is necessary to adapt the integration time when changing the position of the excitation source. It will most likely be necessary to make multiple measurements with different integration times for every source position in order to utilize the detector's dynamic range efficiently and to avoid overexposure of the sensors. Thus, a longer time is needed for the acquisition of the full data set compared to the previous setup with a constant integration time.

The noise model (N3) is somehow the best case and thus the opposite to (N1). It states that the measurements have a constant coefficient of variation ($\mathrm{CV} = \frac{\sigma_{ij}}{\mathcal{M}_{ij}} = \text{const}$). In practical settings such a noise model would be difficult to achieve. However, it is included here to demonstrate what could be reconstructed in an ideal setting.

The resulting images for the same simulation model as described in section 4.6, using the noise models (N1)–(N3) and noise fractions of 0.5 %–5 %, are shown in figure 5.1. The second stopping criterion based on the chi-square distribution (equation (4.18)) was used in the reconstruction algorithm.

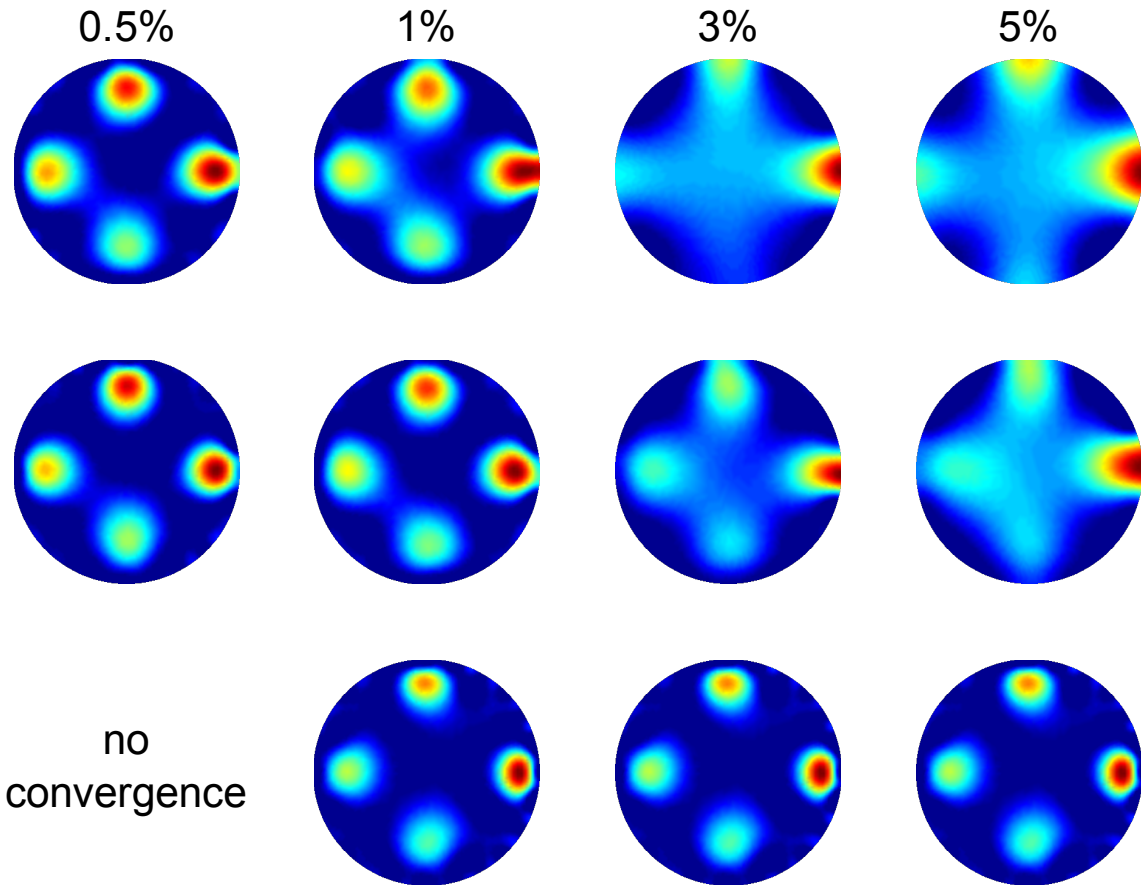

Figure 5.1: Reconstructions using the H^1-semi norm for different noise levels of 0.5 %, 1 %, 3 % and 5 % (columns from left to right) and different noise models based on the largest measurement datum (top row), on the largest measurement per excitation (middle row) and with a constant coefficient of variation bottom row). The stopping criterion was not reached for the low-noise case with the last noise model.

Using the most rigorous noise model (N1), the original fluorophore distribution can be reconstructed sufficiently well up to approximately 1 % Gaußian noise. Already for 3 % noise the individual inclusions are not separated anymore in the reconstruction.

If it is possible to reduce the noise such that it is only dependent on the largest measurement of each excitation (model N2), the results are much clearer. For 0.5 % and 1 % noise the inclusions are better localised and even for a noise fraction of 3 % the four inclusions are recognisable. Only at higher noise levels the reconstructions smear out and become indistinguishable.

For the best noise case (N3), where the coefficient of variation is constant, even the 5 % of data noise do not lead to a corrupted reconstruction. However, there is a problem with the stopping criterion for the low-noise case with a fraction of only 0.5 %. By inspecting the concentration after every update, it is clear that the reconstruction itself works but the variance-weighted residual does not reach the predefined threshold. The reason is not completely clear up to now and needs to be further investigated.

6 Experimental design

I love fools' experiments. I am always making them.

(Charles Darwin)

The placement of sensor and detector optodes is an important parameter for the design of fluorescence tomography hardware. The chosen configuration determines in which region the setup will be sensitive to the fluorophore signal and what the resolution there will be.

The quality of reconstruction usually increases with the amount of sources and detectors in the measurement setup. From a certain number on, however, the measurements will no longer be independent, which is a consequence of the diffusive propagation of light in tissue. This limits the number of optodes in a measurement setup in a natural way. As was shown in chapter 3, also the time needed to solve the inverse problem is directly related to the number of sources and detectors in the measurement setup. Furthermore, each additional optode makes the hardware more expensive to set up and to maintain. Therefore, the question arises on how to position as few optodes as possible in such a manner that the maximum information is gained from a measurement.

In previous work, Graves *et al* [14] have investigated the influences of the number of sources and detectors and their respective distances on the reconstructions. The comparison is based on a singular value analysis of the weight matrix which is essentially the discretised sensitivity matrix but was actually generated from a Born approximation. This method was later on extended by Lasser *et al* [15], who applied it to 360° projection tomography. The drawback of this methodology is that only full measurement setups can be compared as the singular value decomposition operates on the whole sensitivity matrix.

This chapter is based on material published in [44] with written permission by SPIE.

The method presented in the following relaxes these constraints and allows for the computation of a quality criterion for every single optode, i.e. one obtains an importance measure for every optode in the system. Additionally focus regions can be specified to which the measurement setup will be more sensitive to than the rest of the domain.

6.1 Redundancy minimisation

The optimisation method is based on the technique proposed by Curtis *et al* [45]. The idea is to have only independent measurements, such that every measurement results in information which is useful for the reconstruction step. One possibility to quantify the independence between the k-th and l-th measurements is through the computation of the inner-product of the corresponding sensitivities $F'_k(c)$ and $F'_l(c)$. In the following, we denote the discretised derivative $F'(c)$ by J (which refers to the *Jacobian*) in favour of a concise mathematical notation. Then, the independence can be quantified by the inner product of the row-vectors j_k and j_l.

In section 3.3 it was shown that one has to solve one PDE for every source and detector, i.e. there is no use in dismissing single measurements as this would reduce the computational burden only a little bit: the application of F' and F'^* would become slightly cheaper but setting up the matrices is nearly as expensive as for the full matrix. Furthermore, excluding a single measurement does not reduce hardware cost because still the same number of optodes is required.

Therefore, lumped quality measures for optodes rather than single measurements are necessary. In order to arrive at a compact formulation we enumerate the measurements (i.e. the source/detector pairs) such that each measurement is identified uniquely by an index. Let $\mathcal{I}$ be this set of measurement indices. The size of this set is $|\mathcal{I}|$ which is also the number of rows of the sensitivity matrix J. The sub-set of measurements belonging to the i-th source is denoted by $\mathcal{S}_i \subset \mathcal{I}$. By re-ordering the measurements m it is possible to partition them into two sets m_1, m_2 where m_2 consists of all measurements made with the i-th source and m_1 contains all other measurements. The number of rows of the corresponding sub-matrices J_1,J_2

is $|\mathcal{I} \setminus \mathcal{S}_i|$ for J_1 and $|\mathcal{S}_i|$ for J_2. A differential measurement can be written as

$$J = \begin{pmatrix} J_1 \\ J_2 \end{pmatrix}, \quad m = \begin{pmatrix} m_1 \\ m_2 \end{pmatrix}, \quad \begin{pmatrix} m_1 \\ m_2 \end{pmatrix} = \begin{pmatrix} J_1 \\ J_2 \end{pmatrix} c \tag{6.1}$$

The independence of one measurement m_k made with the i-th source and another measurement m_l made with some other source is given by

$$\frac{(j_k, j_l)^2}{\|j_k\|^2 \, \|j_l\|^2}, \quad k \in \mathcal{S}_i, \quad l \in \mathcal{I} \setminus \mathcal{S}_i, \tag{6.2}$$

This quantity is a real number from the interval $[0, 1]$. If the measurements are orthogonal, it will be zero and it will be close to 1 if the measurements are dependent. In contrast to [45], the square of the cosine of the measurement vectors is used here which is advantageous later on when this criterion is compared to the mutual information measure.

The average redundancy of the i-th source optode is obtained when the inner-products of all measurement vectors belonging to this source with all other measurement vectors are computed. This leads to the expression

$$r_i := \frac{1}{|\mathcal{I} \setminus \mathcal{S}_i| \, |\mathcal{S}_i|} \sum_{l \in \mathcal{I} \setminus \mathcal{S}_i} \sum_{k \in \mathcal{S}_i} \frac{(j_k, j_l)^2}{\|j_k\|^2 \, \|j_l\|^2}, \qquad i = 1, \ldots, ns, \tag{6.3}$$

which is again a quantity from $[0, 1]$. In the final algorithm the quantity

$$q_i := 1 - r_i, \qquad i = 1, \ldots, ns, \tag{6.4}$$

is used as a quality measure for the i-th source. One obtains a quality measure for the k-th detector when the set of source indices $\mathcal{S}_i$ in equation (6.3) is replaced by the set of detectors indices $\mathcal{D}_j$.

The implemented algorithm starts from a pool of feasible source and detector optodes. Having computed the quality measures q_i, $i = 1, \ldots, ns$, for all sources, it remains to find the set of the best sources with respect to a certain stopping criterion. For the sake of simplicity, we assume that there is a certain budget for source optodes, i.e. we would like to find a predefined number of "best" sources. However, a full search amongst the initial set is a combinatorial problem and the time needed

to find the exact solution grows exponentially with the size of the set. Instead the algorithm iteratively dismisses the one source exhibiting the worst quality measure from the pool of optodes. After each iteration the quality measures have to be recomputed again, because excluding an optode most likely increases the independence of all others. The final sources left in the pool are considered to be the best source optodes.

In a second step, the algorithm starts from the full initial pool again but iterates on the detector optodes using the same approach as above. The combination of the best sources and best detectors is taken as a final measurement setup.

The lumped quality measure for an optode in equation (6.3) was defined by an arithmetic average of the quality measure of a single measurement. Thus, one could also think about other averages, for example based on the geometric mean which leads to

$$q_{g,i} := \left[\prod_{n \in \mathcal{I} \setminus \mathcal{S}_i} \left(1 - \frac{1}{|\mathcal{S}_i|} \sum_{m \in \mathcal{S}_i} \frac{(j_m, j_n)^2}{\|j_m\|^2 \, \|j_n\|^2} \right) \right]^{\frac{1}{|\mathcal{I} \setminus \mathcal{S}_i|}}, \qquad i = 1, \ldots, ns. \tag{6.5}$$

As will be shown in 6.3, this version has a closer relation to entropy optimisation methods.

6.2 Focusing

When looking at a part of a sample only, say an organ, it might be desired to bias the placement of the optodes such that in the volume of interest a higher sensitivity is obtained. A straightforward approach is to define a mask vector on the finite element mesh having the structure

$$f := (f_1, \ldots, f_{nt}), \qquad 0 \leq f_i \leq 1, \qquad i = 1, \ldots, nt, \tag{6.6}$$

with nt being the number of elements in the mesh. In the simplest case this vector is binary having a value of 1 inside the volume of interest and 0 everywhere else. For more fine-grained focusing the vector elements can take intermediate values as well.

Instead of operating on the sensitivity matrix J, the optimisation algorithm described above operates on the focused Jacobian $J_F = J\,\mathrm{diag}(f)$.

6.3 Comparison to entropy methods

Optimisation techniques which maximise the entropy have a long history in image reconstruction and have been applied successfully to various tomographic imaging modalities (e.g. [46, 47, 48, 49, 50]). The key idea is that the unknown parameter is seen as a realisation of a random variable having a certain probability density. Then, one seeks to reconstruct that parameter distribution leading to the maximum entropy, for example.

A way to quantify independence in the measurements is via the mutual information (MI). Lets start again from the split data set in equation (6.1). To compute the entropy, the concentration vector c is interpreted as random variable. Additionally the a-priori knowledge of the probability distribution is required. For the sake of easiness, the parameters are assumed to be independent, identically distributed with a normal distribution with a variance of σ_c. The resultant model covariance matrix is then diagonal, i.e. $\mathrm{cov}(c) = \sigma_c I$. The data covariance matrix is obtained by the relation

$$\mathrm{cov}(m) = \mathrm{cov}(Jc) = J\,\mathrm{cov}(c)J^\top = \sigma_c J J^\top, \tag{6.7}$$

or when the split data set is used:

$$\mathrm{cov}\begin{pmatrix} m_1 \\ m_2 \end{pmatrix} = \sigma_c \begin{pmatrix} J_1 J_1^\top & J_1 J_2^\top \\ J_2 J_1^\top & J_2 J_2^\top \end{pmatrix}. \tag{6.8}$$

The entropy $\mathrm{H}(m)$ is a measure for the uncertainty of the data. By using the multivariate normal distribution [51] it is expressed as

$$\mathrm{H}(m) = \tfrac{1}{2} \log\left[(2\pi e)^M \det(\mathrm{cov}(m))\right]. \tag{6.9}$$

If one measures only part of the data, say m_1, the entropy is accordingly

$$\mathrm{H}(m_1) = \tfrac{1}{2} \log \left[(2\pi e)^{M-S_i} \det(\mathrm{cov}(m_1)) \right] \tag{6.10}$$

and the conditional entropy of m_1, given that the measurements m_2 have already been acquired, is

$$\mathrm{H}(m_1|m_2) = \tfrac{1}{2} \log \left[(2\pi e)^{M-S_i} \det(\mathrm{cov}(m_1|m_2)) \right], \tag{6.11}$$

where $\mathrm{cov}(m_1|m_2)$ is the conditional covariance. It can be calculated from equation (6.8) by the Schur complement of $J_2 J_2^\top$ in $JJ^\top$ [52]:

$$\mathrm{cov}(m_1|m_2) = \sigma_c \left(J_1 J_1^\top - J_1 J_2^\top (J_2 J_2^\top)^{-1} J_2 J_1^\top \right). \tag{6.12}$$

The covariance is well defined if $J_2 J_2^\top$ is invertible which is the case if the individual sensitivities, i.e. the rows of J, are linearly independent. If the pseudo-inverse $J_2^\top (J_2 J_2^\top)^{-1}$ was the identity matrix, the conditional covariance would be zero which is plausible because then the concentrations could be exactly reconstructed from the knowledge of m_2 only, i.e. measuring m_1 would be futile.

A low conditional entropy $\mathrm{H}(m_1|m_2)$ means that the uncertainty in the data m_1 decreases significantly if the set m_2 is also measured. Another way to quantify the reduction of the uncertainty in m_1 if m_2 is known beforehand, is through the mutual information

$$\mathrm{MI}(m_1, m_2) := \mathrm{H}(m_1) - \mathrm{H}(m_1|m_2). \tag{6.13}$$

The source tested for removal shall correspond to the set m_2. It can be removed from the pool of optodes safely, if the mutual information to all other measurements is high because then the information gained with this source is also present to a large extent in the set m_1 which was obtained without the current source.

The negative mutual information combined with equations (6.10)–(6.12) leads to

$$-\mathrm{MI}(m_1, m_2) = \tfrac{1}{2} \log \left[\det \left(J_1 J_1^\top - J_1 J_2^\top (J_2 J_2^\top)^{-1} J_2 J_1^\top \right) / \det(J_1 J_1^\top) \right]. \tag{6.14}$$

Again $J_1 J_1^\top$ is invertible due to the linear independent rows of J. Now, it is assumed that this symmetric matrix can be decomposed such that $J_1 J_1^\top = A A^\top$ where the matrix A shall be invertible. A is then the square root of the matrix $J_1 J_1^\top$. The

determinant of the latter is then the product of $\det(A)$, i.e.

$$\det{}^{-1}(J_1 J_1^\top) = \det\left((J_1 J_1^\top)^{-\frac{1}{2}}\right) \det\left((J_1 J_1^\top)^{-\frac{1}{2}}\right). \tag{6.15}$$

Finally, we obtain for the mutual information the expression

$$-\operatorname{MI}(m_1, m_2) = \tfrac{1}{2}\log\left[\det\left(I - (J_1 J_1^\top)^{-\frac{1}{2}} J_1 J_2^\top (J_2 J_2^\top)^{-1} J_2 J_1^\top (J_1 J_1^\top)^{-\frac{1}{2}}\right)\right]. \tag{6.16}$$

A diagonal matrix containing the norms of the row vectors of J is denoted by

$$\operatorname{diag}^{-\frac{1}{2}}\left(JJ^\top\right) = \begin{pmatrix} \frac{1}{\|j_1\|} & & 0 \\ & \ddots & \\ 0 & & \frac{1}{\|j_{|\mathcal{I}|}\|} \end{pmatrix}. \tag{6.17}$$

With this expression, the redundancy from equation (6.3) can also be written as

$$r_i = \frac{1}{|\mathcal{I} \setminus \mathcal{S}_i|\,|\mathcal{S}_i|} \sum_{l=1}^{|\mathcal{I}\setminus\mathcal{S}_i|} \sum_{k=1}^{|\mathcal{S}_i|} \left(\operatorname{diag}^{-\frac{1}{2}}\left(J_1 J_1^\top\right) J_1 J_2^\top \operatorname{diag}^{-\frac{1}{2}}\left(J_2 J_2^\top\right)\right)_{l,k}^2. \tag{6.18}$$

which is just the Frobenius norm of the matrix

$$\operatorname{diag}^{-\frac{1}{2}}\left(J_1 J_1^\top\right) J_1 J_2^\top \operatorname{diag}^{-\frac{1}{2}}\left(J_2 J_2^\top\right), \tag{6.19}$$

and can therefore be written in terms of the trace of this matrix multiplied with its own transpose resulting in

$$r_i = \frac{1}{|\mathcal{I} \setminus \mathcal{S}_i|\,|\mathcal{S}_i|} \operatorname{tr}\left(\operatorname{diag}^{-\frac{1}{2}}\left(J_1 J_1^\top\right) J_1 J_2^\top \operatorname{diag}^{-1}(J_2 J_2^\top) J_2 J_1^\top \operatorname{diag}^{-\frac{1}{2}}\left(J_1 J_1^\top\right)\right). \tag{6.20}$$

The matrix notation for the quality measure $q_i = 1 - r_i$ is then

$$q_i = \frac{1}{|\mathcal{I} \setminus \mathcal{S}_i|} \operatorname{tr}\left(I - \frac{1}{|\mathcal{S}_i|} \operatorname{diag}^{-\frac{1}{2}}\left(J_1 J_1^\top\right) J_1 J_2^\top \operatorname{diag}^{-1}(J_2 J_2^\top) J_2 J_1^\top \operatorname{diag}^{-\frac{1}{2}}\left(J_1 J_1^\top\right)\right). \tag{6.21}$$

In a similar manner, one can derive a matrix notation for the geometric quality

measure defined by equation (6.5). It has the form

$$q_{g,i} = \left[\det \operatorname{diag}\left(I - \frac{1}{|\mathcal{S}_i|}\ \operatorname{diag}^{-\frac{1}{2}}\left(J_1 J_1^\top\right) J_1 J_2^\top \right.\right. \\ \left.\left. \times \operatorname{diag}^{-1}(J_2 J_2^\top) J_2 J_1^\top \operatorname{diag}^{-\frac{1}{2}}\left(J_1 J_1^\top\right)\right)\right]^{\frac{1}{|\mathcal{I}\setminus\mathcal{S}_i|}}. \quad (6.22)$$

Comparing the results of (6.16), (6.21) and (6.22), interesting similarities in the structure of the equations can be found although the expressions are not equal. While the mutual information measure combines the whole conditional covariance matrix into one quality measure through the application of the determinant, the other two equations operate on the diagonal part of this matrix only, i.e. they consider only the variances of the parameter but not the covariances. This means that the mutual information is expensive to compute (it requires matrix inversions and computations of determinants) whereas the redundancy measures are comparatively cheap as the inversion and the calculation of determinants have to be carried out for diagonal matrices only.

A disadvantage of the formulation based on Curtis' approach (equation (6.21)) is that there is no strong relationship between the trace of the covariance matrix and its determinant, i.e. it can be that the trace between two setups increases while the determinant decreases.

This drawback is mitigated to a certain extent when the geometric averaging (equation (6.22)) is used. From the Cauchy-Schwarz inequality $\operatorname{cov}^2(x, y) \leq \operatorname{var}(x) \operatorname{var}(y)$, it follows that the covariance has to decrease if the corresponding variances decrease. Thus, one can argue that a decrease in the quality measure based on the geometric mean $q_{g,i}$ will lead to a decrease in $(-\,\mathrm{MI})$ and therefore increase the mutual information. Thus, an optode which is redundant (low $q_{g,i}$) is likely to exhibit a high mutual information content MI between its own measurements and all other measurements taken with other optodes.

6.4 Results

The adaptation algorithm was tested on a homogeneous cylinder with a diameter of 30 mm and a height of 90 mm (figure 6.1 (a)). The same scattering and absorption parameters as in chapter 2 (table 2.1) were used.

To create an initial optode pool, 48 sources and 48 detectors were placed in a zig-zag pattern on a regular grid around the cylinder (see figure 6.1(a)). The six rings of 16 optodes each were spaced by 10 mm. The stopping criterion was set to eight sources and the same number of detectors, i.e. a maximum budget for the optodes is prescribed.

Three different target regions A, B and C were chosen in order to test the focusing ability of the adaptation algorithm. Their masks are defined as

$$f_A(x,y,z) = \begin{cases} 1 & \text{if } 10 \leq z \leq 20, \\ 0 & \text{else,} \end{cases} \tag{6.23}$$

$$f_B(x,y,z) = \begin{cases} 1 & \text{if } -7.5 \leq z \leq 7.5, \\ 0 & \text{else,} \end{cases} \tag{6.24}$$

and

$$f_C(x,y,z) = \begin{cases} 1 & \text{if } x > 0 \text{ and } -20 \leq z \leq -10, \\ 0 & \text{else,} \end{cases} \tag{6.25}$$

i.e. the target regions are two cylinder slices with different thicknesses and one half-slice.

The outcome of the adaptation procedure for the chosen volumes of interest are depicted in figures 6.1(b)–(d). When focusing on region A, the adaptation procedure delivers the ring of optodes covered by this region which is the intuitive arrangement. For the larger volume B, the best arrangement is symmetric for each ring with respect to the cylinder's axis but not symmetric to its mid-plane. The most interesting result is probably the adaptation to the half-cylinder slice shown in figure 6.1(d). The final optodes agglomerate near the focus region because the sensitivity is highest near the optodes. Interestingly the algorithm also suggests to put some of the optodes behind

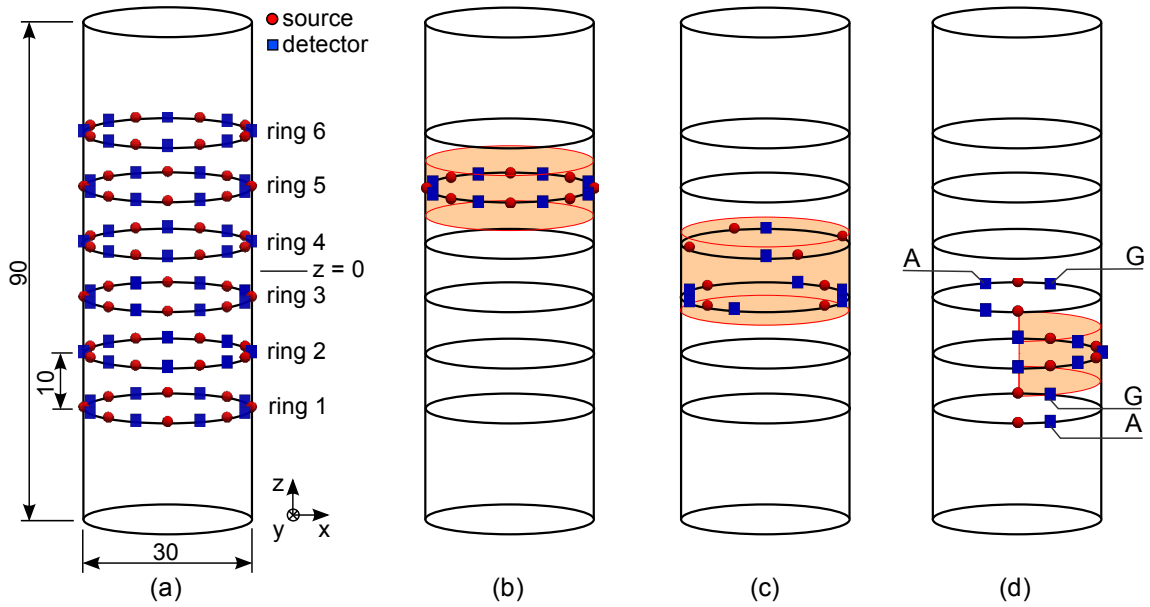

Figure 6.1: (a) Geometry of the optimisation model (measures in mm) together with the initial pool of feasible optodes which are arranged in a zig-zag pattern on six rings with 10 mm spacing. (b)–(d) The result of the adaptation algorithm when focusing on the regions drawn in orange. In the last figure the detectors belonging to the best set found by the geometric or the arithmetic averaging method are marked with G and A, respectively.

the focus region. These optodes measure in combination with an optode at the front fluorescent light across the focus volume and thus provide valuable information.

The stability of the adaptation procedure was tested through a Monte-Carlo simulation. The initial optode locations were shifted randomly with a uniform distribution from $[-1\,mm, 1\text{mm}]$ along the z-axis and the cylinder's perimeter. Thus, the error is up to 10 % of the distance in the z-direction and up to 15 % of the distance along the perimeter. The result of the 50 Monte-Carlo trials is visualised in figure 6.2. The size of the circles (sources) and squares (detectors) is proportional to the frequency of the optode appearing in the best measurement set. While the best set for the half-slice is quite stable as shown in 6.2(b), this is not the case if the full cylinder slice is chosen as target volume (figure 6.2(b)). In the latter, all setups which rotate the optode arrangement by one position around the cylinder axis are equivalent.

The ability to suppress fluorophore signals whose origin is outside the specified target volume is demonstrated by a reconstruction of simulated data from three spherical fluorescent inclusions with a diameter of 5 mm which are located in the focus volumes

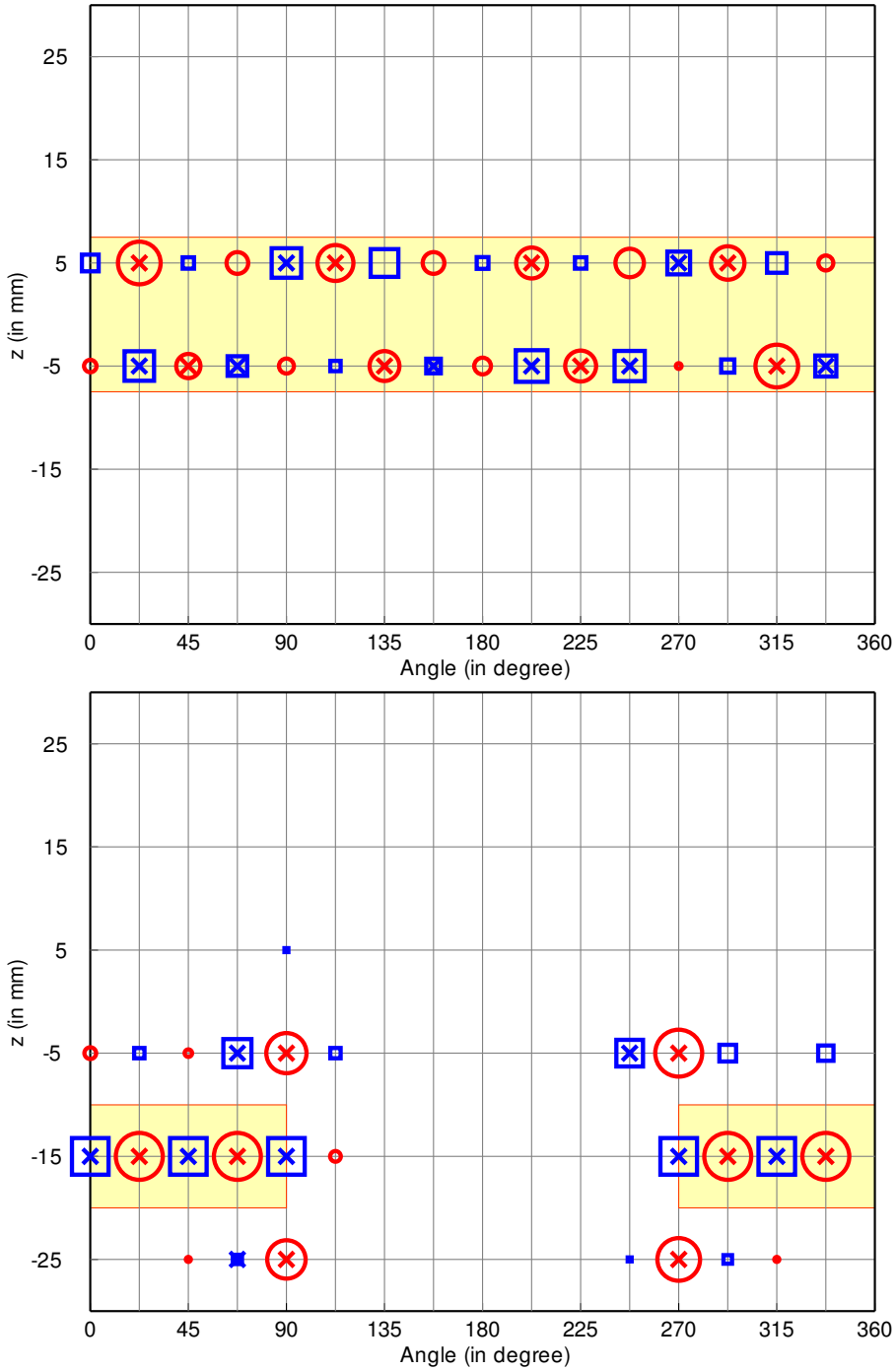

Figure 6.2: Stability of the adaptation result using geometric averaging for focus regions B (a) and C (b). Sources are drawn as circles, detectors as squares. The marker size is proportional to the stability of the optode. Optodes from the reference set (unshifted initial pool of optodes) are marked with "×".

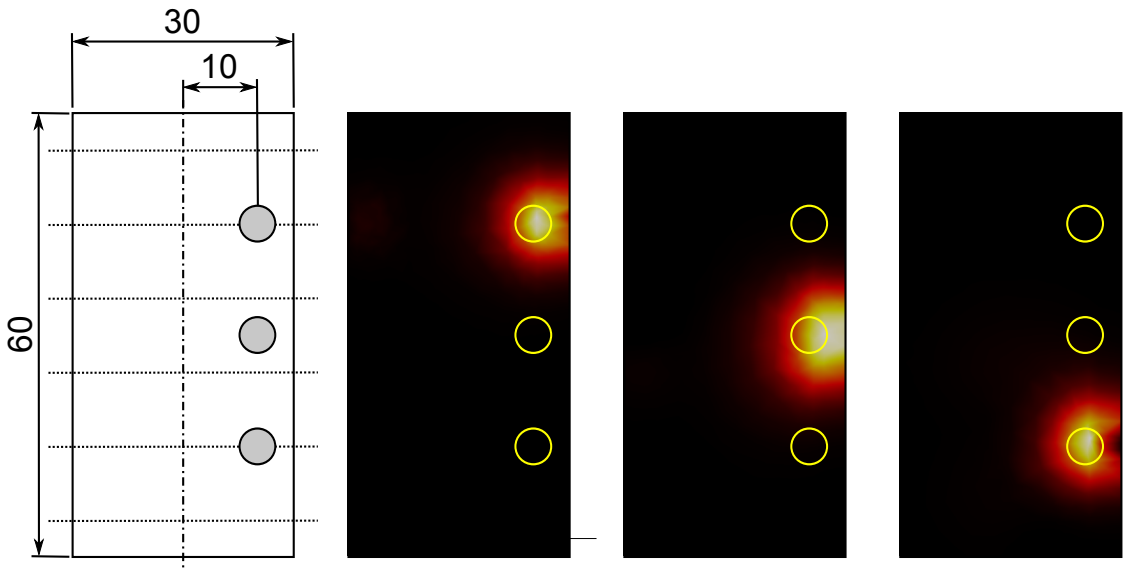

Figure 6.3: (a) The simulation phantom for the generation of measurement data for the reconstruction. The dotted lines mark the height of the six optode rings. The reconstructed fluorophore distribution for three differently focused setups with focus on region A, B and C (from left to right).

according to figure 6.3(a). The optical tissue parameters are again the same as given in table 2.1. The resultant images of the fluorophore distribution using the focused sensitivities are displayed in figure 6.3(b)–(d). In all three cases the off-focus signal is suppressed sufficiently well. The quality of the reconstruction, however, is best for the lower sphere because in this setup the optodes are well concentrated due to the small focus volume.

6.5 Discussion

Contrary to methods previously published by Graves[14] and Lasser [15], the method described above does not rely on a singular value decomposition (SVD) of the sensitivity matrix. As the SVD is a global operation on the Jacobian, it delivers quality measures for whole setups only. The redundancy minimisation approach measures the quality of every optode and therefore provides a much more fine grained information, which might be useful for the design of optimally suited hardware.

Furthermore, the SVD is a comparatively expensive operation from a numerical

point of view. Thus, the method is restricted to 2D geometries or simple 3D geometries with a lower number of finite elements. The method developed by Curtis is again advantageous in this respect as it only relies on the computation of inner products of the Jacobian, which is a cheap operation and can even be parallelised easily, which is not the case for the singular value decomposition.

The stopping criterion used for the results presented in the previous section is just the maximum number of optodes. In practise this would mean that a certain budget can be spent for the sources and detectors. One could also think about other stopping criteria based on the resolution or contrast-to-noise ratio, for example. However, such criteria have not been implemented so far and are subject to further studies.

The resultant algorithm starts from an initial optode pool from which iteratively the worst sensors and detectors are removed. This approach has two drawbacks. The first one is that there is no guarantee that the overall best optodes remain in the pool. Principally, it could be better for the final arrangement, if in a certain iteration the worst optode is kept and the second worst is thrown out, for example. However, the only possibility to make sure of reaching the best set would be an exhaustive search which is exponentially hard to solve with around 6×10^{14} possible solutions for the example presented above. The second drawback is that the optimal set of optodes is not necessarily present in the initial pool. This scenario is actually quite likely considering that there are infinitely many possibilities for the arrangement of a finite amount of optodes on the object's surface.

A viable and promising alternative for both of the previous problems involves formulating the optimisation problem as a distributed control problem with sparsity constraints [53, 54] and to compute a sort of "optode field" that is nonzero only at discrete points. These discrete points can be interpreted as the optimal position for the optode placement. Not only would such an approach circumvent to find the solution of an exponentially growing problem, it would also be independent of an initial optode pool and provide a truly optimal setup, provided that it can be solved by a globally convergent method. Additionally, the optode field's magnitude would be an indicator for the best source strength, i.e. the intensity used for the excitation.

7 Summary

Wenn die anderen glauben,
man ist am Ende, so muss
man erst richtig anfangen.

(Konrad Adenauer)

In this thesis a cross-section of reconstruction and optimization methods for fluorescence diffuse optical tomography was presented. Below, the novelties and shortcomings of these topics shall be summarized.

Nonlinear reconstruction: It was shown that the application of a nonlinear forward model together with non-quadratic regularisation yields more accurate results than a simple linearised reconstruction. Total-variation regularisation and levelset methods have the special advantage that they promote a clear separation of the inclusion from the background and thus do not lead to overly smooth reconstructions.

From a mathematical point of view, the naïve implementation by a smooth relaxation of the total-variation functional and the Heaviside function in the levelset method is not desired. In [55] we could mitigate this problem by using primal-dual methods for solving the inverse problem together with a total-variation regularisation. A mathematically profound version of a levelset reconstruction, based on topological and shape derivatives, is currently under development. First results using only first-order terms of the topological derivative have been published [56].

Optimal optode placement: Surprisingly, the question of how to place the optodes in a fluorescence tomography setup in an optimal manner has yet not received very much attention. The only approaches found in literature restrict themselves to the

comparison of complete setups and are thus only of limited use for optimal hardware design.

With our approach, it is possible to compare individual optodes and to select the best optodes from a predefined pool. Furthermore, we have studied the possibility of suppressing unwanted off-focus signals through biasing the configuration to a chosen target volume.

However, to claim that these positions are optimal would be inappropriate because the optimal optodes are most likely not even part of the initially chosen feasible set. In future work, we intend to overcome this limitation by treating the problem as an optimal control problem, where the sources will be modelled as continuous functions over the boundary of the object. Applying sparsity promoting L^1 constraints to these sources should result in "optode fields" which are nonzero only at discrete points which can then be understood as optimal points.

Furthermore, the experimental evaluation of the methods proposed in this thesis is still pending and subject to further investigation.

8 Acknowledgements

Many thanks go to my supervisor, Prof. Hermann Scharfetter, who initiated this thesis and supported me in every possible way. His unconventional ideas often shed a different light on a topic and helped to break new ground. I am also very grateful to Prof. Barbara Kaltenbacher for her spontaneous acceptance to read and examine the work and pointing out mistakes and unclarities, especially in the mathematical parts.

This thesis would not have been possible without major help from mathematical colleagues. In particular, two people notably influenced this work: Dr. Christian Clason who spent a seemingly endless number of after-work meetings on the explanation of the iteratively regularised Gauß-Newton algorithm and the adjoint system and Dr. Herbert Egger who gave extensive phone support on function spaces, finite element discretisation and preconditioners. Thank you for the cooperation!

I would also like to use this opportunity, to explicitly express my gratitude to all people innervating my live. Unfortunately, I often fail to appreciate their efforts as much as they would deserve. Amongst others, these are Caroline, Isabelle, Andreas, Bernhard, Clemens and Florian.

I am also grateful to my parents and relatives. Without their massive support I most likely would not have started my studies at all.

Last but not least, I would like to thank Bianca for her patience and understanding. At times, I spent more hours with the computer than with her. Thank you for standing by my side when I deserved it least but needed it most!

Bibliography

[1] I. S. Longmuir and J. A. Knopp, "Measurement of tissue oxygen with a fluorescent probe," *J Appl Physiol* **41**(4), pp. 598–602, 1976.

[2] E. Shives, Y. Xu, and H. Jiang, "Fluorescence lifetime tomography of turbid media based on an oxygen-sensitive dye," *Opt. Express* **10**, pp. 1557–1562, 2002.

[3] S. Mordon, V. Maunoury, J. M. Devoisselle, Y. Abbas, and D. Coustaud, "Characterization of tumorous and normal tissue using a ph-sensitive fluorescence indicator (5,6-carboxyfluorescein) in vivo.," *J Photochem Photobiol B* **13**, pp. 307–314, May 1992.

[4] I. Gannot, I. Ron, F. Hekmat, V. Chernomordik, and A. Gandjbakhche, "Functional optical detection based on ph dependent fluorescence lifetime.," *Lasers Surg Med* **35**(5), pp. 342–348, 2004.

[5] Y. Y. Chen and A. W. Wood, "Application of a temperature-dependent fluorescent dye (rhodamine b) to the measurement of radiofrequency radiation-induced temperature changes in biological samples," *Bioelectromagnetics* **30**, pp. 583–590, 2009.

[6] G. Zacharakis, J. Ripoll, R. Weissleder, and V. Ntziachristos, "Fluorescent protein tomograpy scanner for small animal imaging," *IEEE Transactions on Medical Imaging* **24**, pp. 878–885, 2005.

[7] D. S. Elson, I. Munro, J. Requejo-Isidro, J. McGinty, C. Dunsby, N. Galletly, G. W. Stamp, M. A. A. Neil, M. J. Lever, P. A. Kellett, A. Dymoke-Bradshaw, J. Hares, and P. M. W. French, "Real-time time-domain fluorescence lifetime imaging including single-shot acquisition with a segmented optical image intensifier," *New Journal of Physics* **6**, p. 180, 2004.

[8] J. S. Reynolds, T. L. Troy, and E. M. Sevick-Muraca, "Multipixel techniques for frequency-domain photon migration imaging," *Biotechnol. Prog.* **13**, pp. 669–690, 1997.

[9] A. B. Milstein, S. Oh, K. J. Webb, C. A. Bouman, Q. Zhang, D. A. Boas, and R. P. Millane, "Fluorescence optical diffusion tomography," *Applied Optics* **42**, pp. 3081–3094, 2003.

[10] M. J. Eppstein, D. E. Dougherty, T. L. Troy, and E. M. Sevick-Muraca, "Biomedical optical tomography using dynamic parameterization and bayesian conditioning on photon migration measurements," *Applied Optics* **38**, pp. 2138–2150, 1999.

[11] M. A. O'Leary, D. A. Boas, X. D. Li, B. Chance, and Y. G. Yodh, "Fluorescence lifetime imaging in turbid media," *Opt. Lett.* **21**, pp. 158–160, 1996.

[12] V. Ntziachristos and R. Weissleder, "Experimental three-dimensional fluorescence reconstruction of diffuse media by use of a normalized born approximation," *Optics Letters* **26**, pp. 893–895, 2001.

[13] A. Joshi, W. Bangerth, and W. M. Sevick-Muraca, "Adaptive finite element based tomography for fluorescence optical imaging in tissue," *Opt. Express* **12**, pp. 5402–5417, 2004.

[14] E. E. Graves, J. P. Culver, J. Ripoll, R. Weissleder, and V. Ntziachristos, "Singular-value analysis and optimization of experimental parameters in fluorescence molecular tomography," *J. Opt. Soc. Am. A* **21**, pp. 231–241, 2004.

[15] T. Lasser and V. Ntziachristos, "Optimization of 360° projection fluorescence molecular tomography," *Medical Image Analysis* **11**, pp. 389–399, 2007.

[16] L.-H. Wang, S. L. Jacques, and L.-Q. Zheng, "MCML - monte carlo modeling of photon transport in multi-layered tissues," *Computer Methods and Programs in Biomedicine* **47**, pp. 131–146, 1995.

[17] A. P. Gibson, J. C. Hebden, and S. R. Arridge, "Recent advances in diffuse optical imaging," *Physics in Medicine and Biology* **50**, pp. R1–R43, 2005.

[18] S. R. Arridge, "Optical tomography in medical imaging," *Inverse Problems* **15**, pp. R41–R93, 1999.

[19] E. M. Sevick-Muraca, E. Kuwana, A. Godavarty, J. P. Houston, A. B. Thompson, and R. Roy, "Near-infrared fluorescence imaging and spectroscopy in random media and tissue," in *Biomedical Photonics Handbook*, T. Vo-Dinh, ed., ch. 33, CRC Press, Boca Raton, 2003.

[20] H. Egger, M. Freiberger, and M. Schlottbom, "On forward and inverse models in fluorescence diffuse optical tomography," *Inverse Problems and Imaging* **4**(3), pp. 411–427, 2010.

[21] P. Ciarlet, *The finite element method for elliptic problems*, North-Holland, Amsterdam, New York, 1978.

[22] G. Alexandrakis, F. R. Rannou, and A. F. Chatziioannou, "Tomographic bioluminescence imaging by use of a combined optical-pet (opet) system: a computer simulation feasibility study," *Physics in Medicine and Biology* **50**, pp. 4225–4241, 2005.

[23] M. Keijzer, W. M. Star, and P. R. M. Storchi, "Optical diffusion in layered media," *Applied Optics* **27**, pp. 1820–1824, 1988.

[24] A. Joshi, *Adaptive finite element methods for fluorescence enhanced optical tomography*. PhD thesis, Texas A&M University, 2005.

[25] M. L. Landsman, G. Kwant, G. A. Mook, and W. G. Zijlstra, "Light-absorbing properties, stability, and spectral stabilization of indocyanine green," *J Appl Physiol* **40**(4), pp. 575–583, 1976.

[26] A. B. Bakushinskiĭ, "On a convergence problem of the iterative-regularized Gauss-Newton method," *Zh. Vychisl. Mat. i Mat. Fiz.* **32**(9), pp. 1503–1509, 1992.

[27] H. W. Engl, M. Hanke, and A. Neubauer, *Regularization of inverse problems*, vol. 375 of *Mathematics and its Applications*, Kluwer Academic Publishers Group, Dordrecht, 1996.

[28] B. Blaschke, A. Neubauer, and O. Scherzer, "On convergence rates for the iteratively regularized Gauss-Newton method," *IMA J. Numer. Anal.* **17**(3), pp. 421–436, 1997.

[29] A. Tarantola, *Inverse Problem Theory and Methods for Model Parameter Estimation*, SIAM, 1 ed., December 2004.

[30] R. Acar and C. R. Vogel, "Analysis of bounded variation penalty methods for ill-posed problems," *Inverse Problems* **10**, pp. 1217–1229, 1994.

[31] D. Alvarez, P. Medina, and M. Moscoso, "Fluorescence lifetime imaging from time resolved measurements using a shape-based approach," *Optics Express* **17**, pp. 8843–8855, 2009.

[32] M. Schweiger, O. Dorn, and S. R. Arridge, "3-d shape and contrast reconstruction in optical tomography with level sets," *Journal of Physics: Conference Series* **124**(1), p. 012043, 2008.

[33] F. Frühauf, O. Scherzer, and A. Leitão, "Analysis of regularization methods for the solution of ill-posed problems involving discontinuous operators," *SIAM J. Numer. Anal.* **43**, pp. 767–786, 2005.

[34] H. Egger and A. Leitao, "Nonlinear regularization methods for ill-posed problems with piecewise constant or strongly varying solutions," *Inverse Problems* **25**, p. 115014, 2009.

[35] M. Freiberger, H. Egger, and H. Scharfetter, "Nonlinear inversion schemes for fluorescence optical tomography," *IEEE Transactions on Biomedical Engineering* , 2010 (accepted).

[36] H. Dehghani, M. E. Eames, P. K. Yalavarthy, S. C. Davis, S. Srinivasan, C. M. Carpenter, B. W. Pogue, and K. D. Paulsen, "Near infrared optical tomography using nirfast: Algorithms for numerical model and image reconstruction algorithms," *Communications in Numerical Methods in Engineering* , 2009.

[37] V. A. Morozov, "On the solution of functional equations by the method of regularization," *Soviet Math. Dokl.* **7**, pp. 414–417, 1966.

[38] D. Calvetti and E. Somersalo, *Introduction to Bayesian Scientific Computing: Ten Lectures on Subjective Computing (Surveys and Tutorials in the Applied Mathematical Sciences)*, Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2007.

[39] Bortz, *Statistik für Human- und Sozialwissenschaftler (German Edition)*, Springer, 6 ed., September 2005.

[40] W. R. B. Lionheart, "EIT reconstruction algorithms: pitfalls, challenges and recent developments," *Physiological Measurement* **25**(1), p. 125, 2004.

[41] G. Healey and R. Kondepudy, "Radiometric ccd camera calibration and noise estimation," *IEEE Transactions on Pattern Analysis and Machine Intelligence* **16**, pp. 267–276, mar. 1994.

[42] R. Boie and I. Cox, "An analysis of camera noise," *IEEE Transactions on Pattern Analysis and Machine Intelligence* **14**, pp. 671–674, jun. 1992.

[43] Y. Tsin, V. Ramesh, and T. Kanade, "Statistical calibration of ccd imaging process," *Eighth IEEE International Conference on Computer Vision* **1**, pp. 480–487, 2001.

[44] M. Freiberger, C. Clason, and H. Scharfetter, "Adaptation and focusing of optode configurations for fluorescence optical tomography by experimental design methods," *Journal of Biomedical Optics* **15**, pp. 016024–1–016024–10, 2010.

[45] A. Curtis, A. Michelini, D. Leslie, and A. Lomax, "A deterministic algorithm for experimental design applied to tomographic and microseismic monitoring surveys," *Geophysical Journal International* **157**, pp. 595–606, 2004.

[46] A. J. Mwambela, O. Isaksen, and G. A. Johansen, "The use of entropic thresholding methods in reconstruction of capacitance tomography data," *Chemical Engineering Science* **52**, pp. 2149–2159, 1997.

[47] N. V. Denisova, "Maximum-entropy-based tomography for gas and plasma diagnostics," *J. Phys. D: Appl. Phys.* **31**, pp. 1888–1895, 1998.

[48] A. Mohammad-Djafari and G. Demoment, "Maximum entropy image reconstruction in x-ray and diffraction tomography," *IEEE Transactions on Medical Imaging* **7**, pp. 345–354, 1988.

[49] M. K. Nguyen and A. Mohammad-Djafari, "Bayesian approach with the maximum entropy principle in image reconstruction from microwave scattered field data," *IEEE Transactions on Medical Imaging* **13**, pp. 254–262, 1994.

[50] B. A. Ardekani, M. Braun, B. F. Hutton, I. Kanno, and H. Iida, "Minimum cross-entropy reconstruction of pet images using prior anatomical information," *Physics in Medicine and Biology* **41**, pp. 2497–2517, 1996.

[51] N. A. Ahmed and D. V. Gokhale, "Entropy expressions and their estimators for multivariate distributions," *IEEE Transactions on Information Theory* **35**, pp. 688–692, May 1989.

[52] S. Puntanen and G. P. H. Styan, *The Schur Complement and Its Applications (Numerical Methods and Algorithms)*, ch. Schur Complements in Statistics and Probability, pp. 163–226. Springer, 2005.

[53] G. Stadler, "Elliptic optimal control problems with L^1-control cost and applications for the placement of control devices," *Computational Optimization and Applications* , 2007.

[54] C. Clason and K. Kunisch, "A duality-based approach to elliptic control problems in non-reflexive Banach spaces," *ESAIM Control, Optimisation and Calculus of Variations* , 2010.

[55] M. Freiberger, C. Clason, and H. Scharfetter, "Total variation regularization for nonlinear fluorescence tomography with an augmented Lagrangian splitting method," *Applied Optics* **49**, pp. 3741–3747, 2010.

[56] M. Freiberger, M. Hintermüller, A. Laurain, and H. Scharfetter, "Topological derivative for image reconstruction in fluorescence tomography," in *Proceedings 44. Jahrestagung der Deutschen Gesellschaft für Biomedizinische Technik*, 2010.