Modeling and Nonlinear Inversion for
Frequency-Domain Diffuse Optical Tomography

by

Gregory Boverman

B.S., Rensselaer Polytechnic Institute, 1995

A thesis submitted to the
Department of Electrical and Computer Engineering of
Northeastern University
in partial fulfillment of the requirements
for the degree of
Master of Science
2003
This thesis entitled:
Modeling and Nonlinear Inversion for Frequency-Domain Diffuse Optical Tomography
written by Gregory Boveman
has been approved for the Department of Electrical and Computer Engineering

________________________________________
Eric Miller

________________________________________
Prof. Dana Brooks

________________________________________
Prof. Misha Kilmer

________________________________________
Prof. David Boas

Date ________________

The final copy of this thesis has been examined by the signatories, and we find that both the content and the form meet acceptable presentation standards of scholarly work in the above mentioned discipline.
Boverman, Gregory (M.S.E.E.)

Modeling and Nonlinear Inversion for Frequency-Domain Diffuse Optical Tomography

Thesis directed by Prof. Eric Miller

Interest in the tomographic imaging of the body using near infrared (NIR) light has been increasing steadily in recent years. Already, NIR measurement of gross blood oxygenation is a valuable clinical tool, and it is hoped that the tomographic approach will make it possible to dynamically image the spatial distribution of blood oxygenation, in addition to other phenomena that induce optical contrast, with very high temporal resolution.

Before Diffuse Optical Tomography (DOT) can become a useful clinical tool, however, a number of technical challenges must be addressed. First of all, due to the highly scattering nature of light propagation in biological tissue, optical tomography is an inherently three-dimensional problem. Also, the nonlinear relationship between the optical parameters and the detected photon fluence precludes the use of efficient, single-step inversion algorithms possible for other imaging modalities, such as Computed Tomography (CT), and Magnetic Resonance Imaging (MRI).

This thesis focuses on the problem of DOT using sinusoidally modulated sources and frequency-domain measurements. The principal technical contributions of this work are twofold. First, we have developed a finite-difference frequency-domain (FDFD) algorithm to rapidly solve for the three-dimensional photon intensity distribution due to a point source, given the spatial distribution of the optical parameters and the medium geometry. Secondly, we have implemented algorithms to estimate the distribution of optical parameters given measurements of the amplitude and phase at a discrete set of detector locations. To
accomplish this end, we use first and second-order nonlinear optimization methods, with the Jacobian of each measurement computed using the adjoint field method. We also analyze the cost function with respect to the position, size, and absorption contrast of a spherical inclusion, computing statistical bounds on our ability to estimate these parameters. Results are shown for both simulated and phantom measurements.
Dedication

To my parents
Acknowledgements

I thank all the fluffy kitties.
Contents

Chapter

1 Introduction
   1.1 Previous work ........................................... 4
   1.2 Outline of the thesis .................................... 7

2 Forward Models
   2.1 Radiative Transport and the diffusion approximation ........... 8
   2.2 Computational Approaches .................................. 10
   2.3 Finite Difference Modeling .................................. 12
   2.4 Validation of the Finite Difference Model ....................... 16

3 Nonlinear inversion
   3.1 Nonlinear inversion formulation ................................ 26
   3.2 Efficient computation of measurement Jacobians ............... 29
   3.3 Transformation of measured quantities ........................ 37
   3.4 Inversion using the adjoint field method - parametric approach  40
   3.5 Cramér-Rao Lower Bound .................................... 43
   3.6 Noise Modeling ............................................. 45
   3.7 Error Function Analysis ..................................... 49
   3.8 Cramér-Rao bound analysis .................................. 62
4 Inversion results
   4.1 Simulation results ........................................... 67
       4.1.1 Nonparametric Inversion ................................. 67
       4.1.2 Parametric Inversion .................................... 73
   4.2 Parametric inversion in the case of model mismatch ....... 84
   4.3 Phantom Results .............................................. 88

5 Conclusion and Future Work .................................... 95

Bibliography .................................................. 97
Tables

Table

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1</td>
<td>Measurement transformations</td>
<td>39</td>
</tr>
<tr>
<td>3.2</td>
<td>Transformed measurement functional gradients</td>
<td>39</td>
</tr>
<tr>
<td>4.1</td>
<td>Final WMSE and penalty when true object $\mu_a = 0.3 \text{ cm}^{-1}$</td>
<td>72</td>
</tr>
<tr>
<td>4.2</td>
<td>Final WMSE and penalty when true object $\mu_a = 0.2 \text{ cm}^{-1}$</td>
<td>72</td>
</tr>
<tr>
<td>4.3</td>
<td>Final Weighted Mean-Squared Error and parameter estimates for various algorithms</td>
<td>75</td>
</tr>
<tr>
<td>4.4</td>
<td>Final Weighted Mean-Squared Error and parameter estimates for an absorbing sphere with $\mu_a = 0.3 \text{ cm}^{-1}$ for amplitude measurements</td>
<td>77</td>
</tr>
<tr>
<td>4.5</td>
<td>Final Weighted Mean-Squared Error and parameter estimates for an absorbing sphere with $\mu_a = 0.2 \text{ cm}^{-1}$, using amplitude measurements</td>
<td>81</td>
</tr>
<tr>
<td>4.6</td>
<td>Final Weighted Mean-Squared Error and parameter estimates for an absorbing sphere with $\mu_a = 0.1 \text{ cm}^{-1}$, using amplitude measurements</td>
<td>81</td>
</tr>
<tr>
<td>4.7</td>
<td>Final Weighted Mean-Squared Error and parameter estimates for an absorbing sphere with $\mu_a = 0.2 \text{ cm}^{-1}$, amplitude and phase measurements</td>
<td>83</td>
</tr>
</tbody>
</table>
4.8 Final WMSE and penalty for imaging-type absorption reconstructions when the true objects are absorbing ellipsoids of various eccentricities. ......................................................... 87
4.9 Final WMSE and parameter estimates where the true objects are absorbing ellipsoids of various eccentricities. ......................................................... 87
### Figures

#### Figure

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Comparison of the semi-infinite medium analytical solution with</td>
<td>18</td>
</tr>
<tr>
<td></td>
<td>the finite difference solution, for slabs of various widths: magnitude</td>
<td></td>
</tr>
<tr>
<td>2.2</td>
<td>Comparison of the semi-infinite medium analytical solution with</td>
<td>19</td>
</tr>
<tr>
<td></td>
<td>the finite difference solution, for slabs of various widths: phase</td>
<td></td>
</tr>
<tr>
<td>2.3</td>
<td>Comparison of the MC and FD solutions for a homogeneous cube:</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>magnitude</td>
<td></td>
</tr>
<tr>
<td>2.4</td>
<td>Comparison of the MC and FD solutions for a homogeneous cube:</td>
<td>21</td>
</tr>
<tr>
<td></td>
<td>phase</td>
<td></td>
</tr>
<tr>
<td>2.5</td>
<td>Magnitude difference produced by a spherical absorber 1 cm in</td>
<td>22</td>
</tr>
<tr>
<td></td>
<td>diameter, with $\mu_a = 0.1 \text{ cm}^{-1}$, for cubes of various widths</td>
<td></td>
</tr>
<tr>
<td>2.6</td>
<td>Magnitude difference produced by a spherical absorber 1 cm in</td>
<td>23</td>
</tr>
<tr>
<td></td>
<td>diameter, with $\mu_a = 0.1 \text{ cm}^{-1}$, for cubes of various widths</td>
<td></td>
</tr>
<tr>
<td>2.7</td>
<td>Comparison of the magnitude difference, relative to a homogeneous</td>
<td>25</td>
</tr>
<tr>
<td></td>
<td>background, caused by the presence of an absorber 2 cm in diameter, with</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\mu_a = 0.1 \text{ cm}^{-1}$, computed by means of Monte Carlo and</td>
<td></td>
</tr>
<tr>
<td></td>
<td>finite difference simulations.</td>
<td></td>
</tr>
<tr>
<td>3.1</td>
<td>Source-detector geometry and position of true absorbing sphere</td>
<td>50</td>
</tr>
</tbody>
</table>
3.2 The mean-squared magnitude error, for a modulation frequency of 200 MHz, as the estimated \((x, y)\) position of a spherical absorber is varied, with all other parameters assuming their correct values. Also shown are the isocontours of the error function superimposed with vectors in the direction \((\frac{\partial e}{\partial x_o}, \frac{\partial e}{\partial y_o})\).

3.3 The mean-squared magnitude error, for a modulation frequency of 50 MHz, as the estimated \((x, y)\) position of a spherical absorber is varied, with all other parameters assuming their correct values. Also shown are the isocontours of the error function superimposed with vectors in the direction \((\frac{\partial e}{\partial x_o}, \frac{\partial e}{\partial y_o})\).

3.4 The mean-squared phase error, for a modulation frequency of 200 MHz, as the estimated \((x, y)\) position of a spherical absorber is varied, with all other parameters assuming their correct values. Also shown are the isocontours of the error function superimposed with vectors in the direction \((\frac{\partial e}{\partial x_o}, \frac{\partial e}{\partial y_o})\).

3.5 The mean-squared phase error, for a modulation frequency of 50 MHz, as the estimated \((x, y)\) position of a spherical absorber is varied, with all other parameters assuming their correct values. Also shown are the isocontours of the error function superimposed with vectors in the direction \((\frac{\partial e}{\partial x_o}, \frac{\partial e}{\partial y_o})\).

3.6 Isocontours of the mean-squared magnitude error, for a modulation frequency of 200 MHz, as the estimated radius and absorption of a spherical absorber are varied, with all other parameters assuming their correct values. Shown along with the direction of \((\frac{\partial e}{\partial \mu_a}, \frac{\partial e}{\partial c})\) computed using the adjoint-field method.
3.7 Isocontours of the mean-squared magnitude error, for a modulation frequency of 50 MHz, as the estimated radius and absorption of a spherical absorber are varied, with all other parameters assuming their correct values. Shown along with the direction of \( \frac{\partial \mathcal{E}}{\partial \mu_a}, \frac{\partial \mathcal{E}}{\partial \sigma} \) computed using the adjoint-field method. ........................................... 57

3.8 Isocontours of the mean-squared phase error, for a modulation frequency of 200 MHz, as the estimated radius and absorption of a spherical absorber are varied, with all other parameters assuming their correct values. Shown along with the direction of \( \frac{\partial \mathcal{E}}{\partial \mu_a}, \frac{\partial \mathcal{E}}{\partial \sigma} \) computed using the adjoint-field method. ........................................... 58

3.9 Isocontours of the mean-squared phase error, for a modulation frequency of 50 MHz, as the estimated radius and absorption of a spherical absorber are varied, with all other parameters assuming their correct values. Shown along with the direction of \( \frac{\partial \mathcal{E}}{\partial \mu_a}, \frac{\partial \mathcal{E}}{\partial \sigma} \) computed using the adjoint-field method. ........................................... 59

3.10 WMSE, for along the hyperbola: \( \mu_a r^3 = 0.2 \), as a function of object radius. Top left: amplitude measurements at 200 MHz. Top right: phase measurements at 200 MHz. Bottom left: amplitude measurements at 50 MHz. Bottom right: phase measurements at 50 MHz. ................................................................. 61

3.11 The Cramér-Rao lower bound for estimating the x (left) and y (right) position and depth (bottom) of an absorbing sphere, as we vary the sphere radius and absorption ....................................................... 64

3.12 The Cramér-Rao lower bound for estimating the radius (left) and absorption (right) of an absorbing sphere, as we vary the sphere radius and absorption ....................................................... 65
3.13 Uncertainty in the radius estimate, where the true radius is 1 cm, as a function of object absorption, comparing amplitude measurements only with amplitude and phase measurements, for a number of modulation frequencies ........................................ 66

4.1 Source-detector geometry and absorber location for inversion experiments .................................................. 68

4.2 True solution for an absorbing sphere with \( \mu_a = 0.3 \text{cm}^{-1} \) in a homogeneous background ............................................................ 69

4.3 Reconstructed non-parametric solution for an absorbing sphere with \( \mu_a = 0.3 \text{cm}^{-1} \) using amplitude measurements .......................................................... 70

4.4 Reconstructed non-parametric solution for an absorbing sphere with \( \mu_a = 0.2 \text{cm}^{-1} \), using amplitude measurements .......................................................... 71

4.5 Parametric errors, as a function of iteration, for several optimization approaches, using amplitude measurements. .......................................................... 76

4.6 Parametric errors when the true object \( \mu_a = 0.3 \text{cm}^{-1} \), using amplitude measurements .......................................................... 78

4.7 Parametric errors when the true object \( \mu_a = 0.2 \text{cm}^{-1} \), using amplitude measurements .......................................................... 79

4.8 Parametric errors when the true object \( \mu_a = 0.1 \text{cm}^{-1} \), using amplitude measurements .......................................................... 80

4.9 Parametric errors when the true object \( \mu_a = 0.2 \text{cm}^{-1} \), amplitude and phase measurements .......................................................... 82

4.10 Source-detector geometry and positions of ellipsoids with major axis lengths of 2.4 cm, 2.8 cm, 3.2 cm, and 3.6 cm, clockwise, from top left .......................................................... 85
<table>
<thead>
<tr>
<th>Section</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.11</td>
<td>True and estimated absorption for ellipsoid with major axis length 2.4 cm</td>
</tr>
<tr>
<td>4.12</td>
<td>True and estimated absorption for ellipsoid with major axis length 3.6 cm</td>
</tr>
<tr>
<td>4.13</td>
<td>Experimental setup for phantom consisting of a spherical absorber in a homogeneous background</td>
</tr>
<tr>
<td>4.14</td>
<td>Nonlinear absorption reconstruction of phantom using log-magnitude data</td>
</tr>
<tr>
<td>4.15</td>
<td>Nonlinear absorption reconstruction of phantom, with unknown source-detector weights estimated iteratively, using amplitude measurements and a doubly stochastic noise model</td>
</tr>
<tr>
<td>4.16</td>
<td>MSE decrease as a function of iteration using parametric reconstruction</td>
</tr>
<tr>
<td>4.17</td>
<td>Error in estimating the position and radius of an absorbing sphere as a function of iteration</td>
</tr>
</tbody>
</table>
Chapter 1

Introduction

Near-infrared diffusive tomographic imaging is beginning to show promise in a number of clinical application settings, including breast tumor detection and characterization [1] [2] [3], monitoring and detection of strokes [4], and imaging of the hemodynamics of brain function [5] [6]. It is not anticipated that the spatial resolution of this technology will match that of MR or computed tomography (CT), but high temporal resolutions are possible with DOT [7] [8]. In the context of cancer detection, there is evidence that tumors differ from normal tissue in their oxygenation properties [9], and that tumor deoxygenation is highly correlated with cancer severity [10] [11].

This thesis is concerned with the question of how to estimate the location, shape, and contrast of compactly supported absorbing anomalies in turbid media from sparse, frequency-domain boundary measurements of optical flux. The assumption of spatial compactness may be accurate for tumors, which are spatially localized and differ markedly from normal tissue in their optical properties. We formulate the inversion as a nonlinear, iterative optimization problem, using our prior information about the compact support of the anomaly in two ways.

In the first approach, termed “imaging”, we solve for the absorption at every point in space, with the prior information encoded as an edge-preserving penalty term. The addition of this term tends to penalize solutions that are too diffusive,
but allows some edges to remain as compared to a Tikhonov regularizer.

In the second approach, we assume that the anomaly has a particular parametric representation, for the purpose of this thesis that of a sphere, but the general approach extends without difficulty to more complex representations. The two approaches are compared for cases where the model matches the data, and for the case where the actual anomaly is in fact an ellipsoid. For optical tomography we show that the parametric representation is particularly useful because of the great difficulty in discriminating between object size and contrast, demonstrated by analyzing the error surface with respect to these parameters. An additional technical contribution of this thesis is the application of the adjoint field method to analytically compute the gradient of the cost function with respect to the object parameters, as described in Norton [12] for the inverse acoustic problem.

A major difficulty faced by optical tomography is the inherent nonlinearity of the inverse problem, a property that greatly increases its computational difficulty as compared to linear imaging modalities such as MR and CT. The nonlinearity is especially severe for high-contrast anomalies, the case we will examine in this thesis. Essentially, the problem is that photons in a turbid medium do not travel ballistically between the source and detectors, but instead typically experience many scattering events. It is possible to use backprojection-type inversion methods for early-arriving photons [1], which experience few scattering events, but this limits sensitivity considerably. For frequency-domain imaging, the problem domain that we are considering in this thesis, the diffusive imaging inverse problem is to estimate the spatially varying coefficients of a second-order partial differential equation (PDE) from a limited number of boundary measurements. Born and Rytov approximations [13] have been successfully employed to linearize the inverse problem [14], but their validity is limited in the case of high-contrast or large inhomogeneities. Also, linear approximations typically ignore the com-
plexities of the boundary conditions. In the PDE formulation, however, boundary conditions can be dealt with naturally.

Additionally, the inverse problem for optical tomography is extremely ill-posed, meaning that it does not possess a unique solution [15] [16] and it is therefore highly sensitive to measurement noise. The basic reason for this ill-posedness is that optical tomography of biological tissue with light in the 600-800 nm. range is a near-field imaging problem, with features of interest far smaller than the wavelength of the Diffusive Photon Density Waves (DPDW’s), making the detection of high-spatial-frequency variations in image parameters difficult. DOT is generally further complicated by its limited-view nature, as many systems depend on reflectance and transmission measurements in which there is not full angular coverage of the medium being imaged.

In this thesis, we compare two approaches to stabilizing the inverse problem. In the imaging approach, we estimate the absorption at each point in space, regularizing by penalizing the cost function with a term that is a function of the solution. However, regularized imaging suffers from the disadvantage of having a weighting parameter that must be adjusted, either empirically or statistically, by means of cross-validation. A number of methods for adjusting this parameter, mainly justified using heuristic arguments, have been reported in the literature. As examples, we refer to the L-curve and Generalized Cross-Validation [17], but these methods typically involve multiple runs of the inversion algorithm, with a correspondingly high computational cost. In the parametric approach, we assume that the solution of the inverse problem belongs to a particular functional form, which is, in the case of this thesis, that of a spherical absorber in a medium whose background structure is known. This may be an accurate assumption for tumors, which are often spatially compact and contrast sharply from the background in absorption [2].
We show experimentally that the parametric approach offers an alternative to imaging that is robust to measurement noise and that does not require the adjustment of a regularization parameter. We believe that this is a promising initial result, and that the extension of the parametric approach to more complex functional forms offers a highly appealing alternative to Tikhonov-type regularization.

In order to solve the inverse problem, we use iterative steepest descent and conjugate gradient methods, which minimize a quadratic function of the measurement error. In the statistical paradigm, assuming Gaussian uncorrelated noise, we are maximizing the log-likelihood or penalized log-likelihood of the parameters given the model. In order to conduct this multi-dimensional minimization efficiently, it is necessary to be able to compute the Jacobian of the cost function with respect to the parameters. Although the Jacobian can be approximated using finite differences, the adjoint field method allows us to compute this gradient with respect to an arbitrarily large parameter set, resulting in significant computational savings.

1.1 Previous work

In recent years, the research interest in image reconstruction algorithms for DOT (or Optical Diffusion Tomography, as it is sometimes referred to) has been steadily growing. Some authors have applied techniques from other disciplines, such as the filtered backprojections used for X-ray tomography, to the optical domain [18] [19], but deblurring is necessary in order to produce reasonable reconstructions.

In the area of fully nonlinear reconstruction, the theoretical basis for the perturbative approach to image reconstruction for optical tomography is described by Arridge [20] [21], who computes the Jacobian of a number of time and frequency-
domain measurement functions with respect to the spatially varying absorption
and scattering parameters. The Jacobians are computed analytically for several
geometries, and a method for computing the measurement Jacobians with respect
to finite element mesh coefficients is also given.

Huabei Jiang and co-workers at Dartmouth University implemented this ap-
proach, showing the effectiveness of the fully nonlinear method for both simulated
and phantom data [22], using Tikhonov regularization and post-reconstruction
spatial lowpass filtering to stabilize the inverse problem. Optimization with re-
spect to the absorbing and scattering perturbations was accomplished using a
Levenberg-Marquardt method. The simulation results showed a high level of sen-
sitivity to noise in the image reconstruction. In a follow-up paper, Paulsen and
Jiang [23] showed that sensitivity to noise and quantitative reconstruction ac-
curacy could be improved by incorporating total variation regularization, which
tends to reinforce sharp edges in the reconstructed image, into the inversion, again
in the context of inversion with respect to finite-element coefficients.

More recent work, grounded in functional analysis and the calculus of vari-
ations, uses perturbation methods to compute functional gradients of the mea-
surements with respect to variations in absorption and scattering. A relatively
early effort in this direction is the paper by Yao et. al. [24], in which the authors
reconstruct the absorption and scattering parameters of strongly absorbing and
scattering inclusions in a two-dimensional geometry using a Born iterative method.
Essentially, the total expected field for the current estimate of the absorption and
scattering is calculated by a multigrid finite-difference solver, and the perturbation
is computed using the Born approximation. The approach given in this paper is
somewhat hindered by the assumption that the spatial variation in the scattering
coefficient is small, an assumption removed by Ye et. al. [25], at Purdue Uni-
versity, who compute the correct functional derivative with respect to scattering
perturbations. In the same paper, the Purdue group proposes an approximate Fréchet derivative with respect to scattering that does not require numerical approximations of gradients. In a subsequent paper [26], the same group reports on using the computed Fréchet derivative in a Bayesian image reconstruction scheme, with a Markov Random Field prior and shot noise measurements.

In the area of parametric inversion, Kilmer et. al. [27] parametrize the boundary of a two-dimensional inclusion using B-splines, and successfully used a greedy algorithm to optimize for the location of the control points. Kolehmainen and co-workers [28] introduced an algorithm for efficiently computing the Jacobian of frequency-domain measurements with respect to the coefficients of a two-dimensional parametrized boundary, assuming that the optical contrast of the piecewise constant inclusions is known. In simulation results [29], with the boundary parametrized using a Fourier basis, it was shown that the shape of the boundary could be accurately reconstructed, with optimization accomplished by means of the Levenberg-Marquardt algorithm. It should be noted that this work is closely related to Kirsch’s “domain derivative” in the inverse scattering field [30]. Recent work in the problem domains of Emission Computed Tomography and inverse scattering has given expressions for computing the Cramer-Rao bounds of parameterized boundary estimates [31] [32].

Remarkably, similar mathematical models and inversion approaches can be employed in a number of disparate fields. Notably, to a reasonable level of approximation, and for certain classes of problems, the scalar Helmholtz equation governs wave propagation in acoustical media [12], the propagation of electric fields in living tissue [33] and in the earth [34] [35], in addition to describing the physics of Diffuse Photon Density Waves (DPDW’s) in scattering media [36].
1.2 Outline of the thesis

In the thesis that follows, we focus on the problem of localizing and quantitatively characterizing breast tumors, which are assumed to be highly optically absorbing of 700 nm. light [2] [3], due to their deoxygenation. Simulation results are shown for a number of model geometries, and the algorithms are also applied to experimental phantom data, collected at Massachusetts General Hospital, as well.

In Chapter 2, we discuss the forward model and compare it to analytical solutions as well as to Monte Carlo simulations of photon transport. We discuss non-parametric and parametric inversion algorithms in Chapter 3, specifically focusing on efficiently computing the Jacobian of the measurements with respect to our inversion parameters. We also present results analyzing the shape of the cost function as we vary the position, absorption, and size of a spherical absorbing anomaly, in a transmission geometry, for a number of modulation frequencies. We compute the Cramer-Rao bounds on the uncertainty of our parameter estimates, as a function of the object size and contrast. We find that phase information improves our ability to characterize an absorbing object, with measurements of the amplitude and phase at 200 MHz giving superior characterization to measurements at 50 MHz or 0 MHz. Simulation results for various object contrasts and signal-to-noise ratios are presented in Chapter 4, validating the conclusions reached by the cost function analysis. Conclusions and suggestions for future work are presented in Chapter 5.
Chapter 2

Forward Models

2.1 Radiative Transport and the diffusion approximation

The distribution of light intensity in highly scattering media, ignoring polarization, can be very accurately modeled using the radiative transport model, developed by Chandrasekhar [37]. Conceptually, this model is very powerful in that it does not require knowing the small-scale variations in a material’s index of refraction, but rather assumes that photons traversing the medium do not interfere with one another, and that photons experience discrete absorption and direction-altering scattering events, with the probability of these events varying spatially and given by the optical absorption and scattering coefficients, respectively. Unlike Maxwell’s Equations, which are deterministic, the main relation governing photon transport, the Radiative Transport Equation, is a stochastic equation. The light intensity, $L$, at position $\mathbf{r}$, in direction $\mathbf{\hat{\Omega}}$, at time $t$ can be given by the following conservation equation [38]:

$$
\frac{1}{v} \frac{\partial L(\mathbf{r}, \mathbf{\hat{\Omega}}, t)}{\partial t} + \nabla \cdot L(\mathbf{r}, \mathbf{\hat{\Omega}}, t) + \mu_t L(\mathbf{r}, \mathbf{\hat{\Omega}}, t) = \mu_s \int L(\mathbf{r}, \mathbf{\hat{\Omega}'}, t) f(\mathbf{\hat{\Omega}}, \mathbf{\hat{\Omega}'}) d\mathbf{\hat{\Omega}'} + S(\mathbf{r}, \mathbf{\hat{\Omega}}, t)
$$

(2.1)

The function $f(\mathbf{\hat{\Omega}}, \mathbf{\hat{\Omega}'})$ gives the probability of scattering from direction $\mathbf{\hat{\Omega}}$ to direction $\mathbf{\hat{\Omega}'}$. The source distribution is given by $S(\mathbf{r}, \mathbf{\hat{\Omega}}, t)$, and the speed of light is $v$. The space-varying optical parameters are the absorption, $\mu_a$, and the scattering, $\mu_s$, with $\mu_t = \mu_s + \mu_a$. 
As the RTE is computationally expensive to simulate, approximations have been developed to simplify the forward modeling problem. An often-used approach is the $P_N$ approximation [39], which expands the luminance and source as a linear combination of spherical harmonic functions, $Y_{l,m}(\hat{\Omega})$.

$$L(\mathbf{r}, \hat{\Omega}, t) = \sum_{l=0}^{N} \sum_{m=-l}^{l} \left( \frac{2l + 1}{4\pi} \right) \frac{1}{2} \phi_{l,m}(\mathbf{r}, t) Y_{l,m}(\hat{\Omega})$$ (2.2)

$$S(\mathbf{r}, \hat{\Omega}, t) = \sum_{l=0}^{N} \sum_{m=-l}^{l} \left( \frac{2l + 1}{4\pi} \right) \frac{1}{2} q_{l,m}(\mathbf{r}, t) Y_{l,m}(\hat{\Omega})$$ (2.3)

In the $P_1$ approximation, the luminance can be written as the sum of monopole and dipole components:

$$L(\mathbf{r}, \hat{\Omega}, t) = \frac{1}{4\pi} \phi(\mathbf{r}, t) + \frac{3}{4\pi} \mathbf{J}(\mathbf{r}, t) \cdot \hat{\Omega}$$ (2.4)

Here $\phi(\mathbf{r}, t)$ is defined to be the photon density, and $\mathbf{J}(\mathbf{r}, t)$ the photon current:

$$\phi(\mathbf{r}, t) = \int L(\mathbf{r}, \hat{\Omega}, t) d\hat{\Omega}$$ (2.5)

$$\mathbf{J}(\mathbf{r}, t) = \int L(\mathbf{r}, \hat{\Omega}, t) \hat{\Omega} d\hat{\Omega}$$ (2.6)

Assuming that $f(\hat{\Omega}, \hat{\Omega}')$ depends only on the angle between $\hat{\Omega}$ and $\hat{\Omega}'$, and using the recurrence relation for spherical harmonics [38] leads to the following set of equations for the $P_1$ approximation:

$$\left( \mu_a(\mathbf{r}) + \frac{1}{v} \frac{\partial}{\partial t} \right) \phi(\mathbf{r}, t) + \nabla \cdot \mathbf{J}(\mathbf{r}, t) = S_0(\mathbf{r}, t)$$ (2.7)

$$\left( \frac{1}{v} \frac{\partial}{\partial t} + (\mu'_s + \mu_a) \right) \mathbf{J}(\mathbf{r}, t) + \frac{1}{3} \nabla \phi(\mathbf{r}, t) = S_1(\mathbf{r}, t)$$ (2.8)

The source monopole and dipole moments are $S_0(\mathbf{r}, t)$, and $S_1(\mathbf{r}, t)$ respectively, and a reduced scattering coefficient, $\mu'_s$, is introduced, where $\mu'_s = (1-g)\mu_s$, $g$ being the average cosine of the scattering angle.
If the further assumptions are made that the source is isotropic (i.e. $S_1(r, t) = 0$), and that $\frac{\partial \mathbf{J}}{\partial t} = 0$, the diffusion approximation results:

$$- \nabla \cdot D(r) \nabla \phi(r, t) + \mu_a(r) + \frac{1}{v} \frac{\partial \phi(r, t)}{\partial t} = S_0(r, t)$$

(2.9)

In the frequency domain, we have:

$$- \nabla \cdot D(r) \nabla \phi(r, \omega) + (\mu_a(r) + \frac{j \omega}{v})\phi(r, \omega) = S_0(r, \omega)$$

(2.10)

with the diffusion coefficient, $D$, defined to be $\frac{1}{3\mu_e}$. This equation shows that, in a homogeneous medium, the photon intensity, $\phi$, propagates in a wavelike manner, with exponential attenuation and linear phase. It has been experimentally shown that the resulting photon density waves experience refraction and diffraction in a manner similar to that exhibited by electromagnetic waves [40] [41].

At the interface between scattering and non-scattering media, the partial-current boundary condition (a Robin-type condition) very effectively models the interactions between photon density waves and boundaries [42]. A relationship is established between the photon intensity and the current normal to the boundary, as follows:

$$\frac{1}{4}R_\phi \phi(r) - \frac{D(r)}{2}R_j \hat{n} \cdot \nabla \phi(r) = 0$$

(2.11)

where $R_\phi$ and $R_j$ are the Fresnel reflection coefficients for the photon density and current, respectively, and $\hat{n}$ is the direction normal to the boundary.

All of our forward-modeling and inversion results will be shown in the context of the diffusion approximation.

### 2.2 Computational Approaches

A number of methods, with widely varying levels of computational and implementation complexity, have been developed to predict the transport of light in living tissue. Monte-Carlo methods [43] simulate the paths taken by individual
photons as they scatter through an inhomogeneous medium. The simulation is stochastic, with the variance of the resulting fluence distribution dependent on the number of photons used. While the complexity of the resulting computer code is not high, Monte-Carlo methods may require extremely large numbers of photons to estimate the three-dimensional fluence distribution with sufficiently low variance.

Methods based on implementation of the transport equation are of considerable sophistication and are also generally computationally intensive, but may be of use in simulating photon migration in media containing both scattering and non-scattering regions [44].

Computational solution of the diffusion equation generally involves numerical simulation of the diffusion partial differential equation or the associated integral equation, in the time or frequency domains [45] [46]. In the time domain, implicit of explicit time-stepping algorithms can be employed to solve for the photon density at each time instant. In the frequency domain, the approach generally taken is to solve for the steady-state photon density resulting from a sinusoidally modulated input, simultaneously at all points in space. To this end, the space is discretized, and a large, generally sparse, linear system is generated, encoding both the PDE and the boundary conditions. This large system can then be inverted by direct or iterative methods.

The Born and Rytov approximations [14] linearize the forward problem using perturbational methods, writing the total photon density as the sum of incident and scattered components.

\[
\phi_{\text{Born}}(\mathbf{r}) = \phi_{\text{in}}(\mathbf{r}) + \phi_{\text{sc}}(\mathbf{r})
\]  \hspace{1cm} (2.12)

\[
\phi_{\text{Rytov}}(\mathbf{r}) = \exp(\phi_{\text{in}}(\mathbf{r}) + \phi_{\text{sc}}(\mathbf{r}))
\]  \hspace{1cm} (2.13)

The Born approximation can be used to image perturbations in absorption, mak-
ing the assumption that the scattered field is small relative to the incident field, which is the case when the absorbing perturbation is small and of low contrast. Converting the diffusion partial differential equation into an integral equation leads to the following:

\[
\phi_{sc}(\mathbf{r}_s, \mathbf{r}_d) = -\int G(\mathbf{r}_d, \mathbf{r}') \frac{\delta \mu_a(\mathbf{r}')}{D} \phi_{in}(\mathbf{r}_s, \mathbf{r}') d^3 r'
\]  

(2.14)

where \( G(\mathbf{r}_d, \mathbf{r}') \) is the Green function solution for the Helmholtz equation. If we assume further that the Green function for an infinite homogeneous medium can be used, ignoring multiple scattering interactions and the complexities of the boundary, we have:

\[
\phi_{sc}(\mathbf{r}_s, \mathbf{r}_d) = -\int \frac{\exp(j k |\mathbf{r}' - \mathbf{r}_d|)}{4\pi |\mathbf{r}' - \mathbf{r}_d|} \frac{\delta \mu_a(\mathbf{r}')}{D} \phi_{in}(\mathbf{r}_s, \mathbf{r}') d^3 r'
\]  

(2.15)

The Rytov approximation, which has wider applicability than the Born approximation for absorbing perturbations, is derived by assuming that the scattered field is slowly varying. The resulting expression, again making use of the Green function for an infinite homogeneous medium is:

\[
\phi_{sc}(\mathbf{r}_s, \mathbf{r}_d) = -\frac{1}{\exp(\phi_{in}(\mathbf{r}_s, \mathbf{r}_d))} \int \frac{\exp(j k |\mathbf{r}' - \mathbf{r}_d|)}{4\pi |\mathbf{r}' - \mathbf{r}_d|} \frac{\delta \mu_a(\mathbf{r}')}{D} \phi_{in}(\mathbf{r}_s, \mathbf{r}') d^3 r'
\]  

(2.16)

Similar linearizations can be performed for scattering perturbations.

2.3 Finite Difference Modeling

We have developed software to rapidly solve the frequency-domain diffusion equation, (2.10) using the finite-difference method, discretized on a uniform grid. This method has the advantage over finite-elements of simplicity, since there is no need to compute a three-dimensional mesh (a difficult computational problem), but some complications still arise in implementing the boundary conditions.

The finite difference method expands the photon density at each point in a Taylor series, implicitly assuming that the density is continuous and differentiable,
and truncating the series to a specified accuracy. For example, for a uniform
grid, where \( h \) is the grid spacing, we can approximate the photon density at
\( \mathbf{r} = (x + 1, y, z) \) and \( \mathbf{r} = (x - 1, y, z) \) as follows:

\[
\phi^{x+1,y,z} = \phi^{x,y,z} + h \frac{\partial \phi^{x,y,z}}{\partial x} + \frac{h^2}{2} \frac{\partial^2 \phi^{x,y,z}}{\partial x^2} + \frac{h^3}{6} \frac{\partial^3 \phi^{x,y,z}}{\partial x^3} + \frac{h^4}{24} \frac{\partial^4 \phi^{x,y,z}}{\partial x^4} + \cdots \tag{2.17}
\]

\[
\phi^{x-1,y,z} = \phi^{x,y,z} - h \frac{\partial \phi^{x,y,z}}{\partial x} + \frac{h^2}{2} \frac{\partial^2 \phi^{x,y,z}}{\partial x^2} - \frac{h^3}{6} \frac{\partial^3 \phi^{x,y,z}}{\partial x^3} + \frac{h^4}{24} \frac{\partial^4 \phi^{x,y,z}}{\partial x^4} + \cdots \tag{2.18}
\]

Summing these two equations:

\[
\frac{\phi^{x+1,y,z} - 2\phi^{x,y,z} + \phi^{x-1,y,z}}{h^2} = \frac{\partial^2 \phi^{x,y,z}}{\partial x^2} + \frac{h^2}{12} \frac{\partial^4 \phi^{x,y,z}}{\partial x^4} + \cdots \tag{2.19}
\]

Proceeding in this way, a second-order approximation to \( \nabla^2 \phi \), which appears
in the diffusion equation in the case of homogeneous scattering, can be derived:

\[
\nabla^2 \phi^{x,y,z} = \frac{\partial^2 \phi^{x,y,z}}{\partial x^2} + \frac{\partial^2 \phi^{x,y,z}}{\partial y^2} + \frac{\partial^2 \phi^{x,y,z}}{\partial z^2} \tag{2.20}
\]

\[
\frac{\partial^2 \phi^{x,y,z}}{\partial x^2} \approx \frac{\phi^{x+1,y,z} - 2\phi^{x,y,z} + \phi^{x-1,y,z}}{h^2} + O(h^2) \tag{2.21}
\]

\[
\frac{\partial^2 \phi^{x,y,z}}{\partial y^2} \approx \frac{\phi^{x,y+1,z} - 2\phi^{x,y,z} + \phi^{x,y-1,z}}{h^2} + O(h^2) \tag{2.22}
\]

\[
\frac{\partial^2 \phi^{x,y,z}}{\partial z^2} \approx \frac{\phi^{x,y,z+1} - 2\phi^{x,y,z} + \phi^{x,y,z-1}}{h^2} + O(h^2) \tag{2.23}
\]

Along the same lines, similar approximations can be developed when the
grid spacing is not uniform.

Arridge [45] has presented a finite-difference approximation to the diffusion
equation, for the case of a uniform grid. For the purpose of stability, the assumption
is made that the grid for the diffusion coefficient, \( D \), is offset from the grid
for that absorption coefficient, \( \mu_a \), and \( \phi \) by \( \frac{1}{2} h \). In three dimensions, we get the
following approximation for (2.10):

\[
-(\frac{\partial}{\partial x} D^{x,y,z} \frac{\partial \phi^{x,y,z}}{\partial x} + \frac{\partial}{\partial y} D^{x,y,z} \frac{\partial \phi^{x,y,z}}{\partial y} + \frac{\partial}{\partial z} D^{x,y,z} \frac{\partial \phi^{x,y,z}}{\partial z}) + \frac{j \omega}{\nu} \mu_a^{x,y,z} \phi^{x,y,z} = S_0^{x,y,z} \tag{2.24}
\]
\[
\frac{\partial}{\partial x} D^{x,y,z} \frac{\partial \phi^{x,y,z}}{\partial x} \approx \frac{1}{h} \left( \frac{D^{x+\frac{1}{2},y,z} \phi^{x+1,y,z} - \phi^{x,y,z} - \phi^{x-1,y,z}}{h} \right) \quad (2.25)
\]
\[
\frac{\partial}{\partial y} D^{x,y,z} \frac{\partial \phi^{x,y,z}}{\partial y} \approx \frac{1}{h} \left( \frac{D^{x,y+\frac{1}{2},z} \phi^{x,y+1,z} - \phi^{x,y,z} - \phi^{x,y-1,z}}{h} \right) \quad (2.26)
\]
\[
\frac{\partial}{\partial z} D^{x,y,z} \frac{\partial \phi^{x,y,z}}{\partial z} \approx \frac{1}{h} \left( \frac{D^{x,y,z+\frac{1}{2}} \phi^{x,y,z+1} - \phi^{x,y,z} - \phi^{x,y,z-1}}{h} \right) \quad (2.27)
\]

The boundary condition, (2.11), is implemented using first-order differences

\[
\frac{\partial \phi^{x,y,z}}{\partial x} \approx \frac{\phi^{x+1,y,z} - \phi^{x,y,z}}{h} \approx \frac{\phi^{x,y,z} - \phi^{x-1,y,z}}{h} \quad (2.28)
\]
\[
\frac{\partial \phi^{x,y,z}}{\partial y} \approx \frac{\phi^{x,y+1,z} - \phi^{x,y,z}}{h} \approx \frac{\phi^{x,y,z} - \phi^{x,y-1,z}}{h} \quad (2.29)
\]
\[
\frac{\partial \phi^{x,y,z}}{\partial z} \approx \frac{\phi^{x,y,z+1} - \phi^{x,y,z}}{h} \approx \frac{\phi^{x,y,z} - \phi^{x,y,z-1}}{h} \quad (2.30)
\]

where the precise set of equations that are used depends on the points actually present at the boundary.

For the sake of simplicity, the boundary is taken to coincide with the grid points, which is justified by the fact that the diffusion equation is a smoothing operator whose solution tends not to be affected by variations of high spatial frequency. This is, incidentally, the same property that makes the solution of the inverse problem so difficult.

A further difficulty is the estimation of the surface normal vectors, \( \hat{n} \) in (2.11). Our system accomplishes this by estimating the boundary surface locally, at each point, as a bicubic polynomial parametrized by \( u \) and \( v \). The coefficients of \( f_x(u, v) \), \( f_y(u, v) \), and \( f_z(u, v) \) are computed in a least-squares fashion, and the local normal vector can be determined by:

\[
\hat{n} = \left( \frac{\partial f_x(u, v)}{\partial u}, \frac{\partial f_y(u, v)}{\partial u}, \frac{\partial f_z(u, v)}{\partial u} \right) \times \left( \frac{\partial f_x(u, v)}{\partial v}, \frac{\partial f_y(u, v)}{\partial v}, \frac{\partial f_z(u, v)}{\partial v} \right) \quad (2.31)
\]

The finite-difference equations for the internal and boundary points are encoded into a large, sparse linear system, which can then be solved by direct or
iterative methods (or some combination of the two). The resulting system is:

\[
\begin{bmatrix}
A_R & -A_3 \\
A_3 & A_R
\end{bmatrix}
\begin{bmatrix}
\Phi_R \\
\Phi_3
\end{bmatrix} =
\begin{bmatrix}
S_R \\
S_3
\end{bmatrix}
\] (2.32)

where where \(A_R\) contains the real part of the discrete approximation to the diffusion operator, and \(A_3\) contains its imaginary part, which is \(\frac{\pi}{6}I\). The photon density, \(\Phi\) and sources, \(S\) are also decomposed into their real and imaginary components. We have chosen to index our system such that boundary points receive low indices, followed by the internal points. Thus, \(A_R\) can be further decomposed as follows:

\[
A_R = \begin{bmatrix}
B & BN \\
NB & N
\end{bmatrix}
\] (2.33)

where \(B\) is diagonal, representing our constraint that boundary points cannot depend on other boundary points, \(N\) is symmetric positive-definite and \(NB \neq BN\). The system, (2.32), is extremely large, possibly containing hundreds of thousands of rows for realistic three-dimensional simulations, and it is also extremely sparse, with at most seven non-zero elements per row, for internal points. The system is therefore highly amenable to solution by iterative methods, which, rather than attempting to solve the system directly, minimize the residual, \(\|A\Phi - S\|_2\). Although the minimization of this residual does not guarantee the minimization of the error \(\|\Phi - A^{-1}S\|_2\), the two should coincide for a reasonably well-conditioned \(A\).

A great deal of research has focused on solving a large linear system by projection onto a Krylov subspace \(v_0, Av_0, A^2v_0, A^3v_0, \cdots\), where \(v_0 = \frac{S - A\Phi_0}{\|S - A\Phi_0\|}\), \(\Phi_0\) being an initial guess for the solution. In the case where \(A\) is symmetric and positive-definite, the conjugate gradient algorithm can compute this projection efficiently using a short-term recursion [47]. Recently, algorithms have been developed to solve large non-symmetric systems by projection onto Krylov sub-
spaces, making use of the Arnoldi algorithm to generate an $A$-orthogonal basis for the Krylov subspace [48]. To solve the frequency-domain diffusion equation, we make use of the GMRES algorithm, with a restart chosen experimentally [49].

Convergence of the iterative method can be improved by means of a preconditioner. We have implemented left preconditioning, in which case the residual minimized is $\|M^{-1}A\Phi - M^{-1}S\|_2$. Typically, $M$ is an approximation to $A$ that can be inverted with little computational effort. We have made use of three fairly elementary preconditioners, Jacobi, SSOR(Symmetric Successive Over-Relaxation, and ILU(0). If $A$ is decomposed as $D - E - F$, $F$ and $E$ being the upper and lower-triangular parts of $A$, respectively,

$$M_J = D$$

$$M_{SSOR} = (D - \omega E)D^{-1}(D - \omega F)$$

ILU(0) is an incomplete LU factorization of $A$, without fill-in. The Jacobi and SSOR preconditioners require storage of $n$ elements, $n$ being the number of rows in $A$, while ILU(0) requires storage equal to that of $A$. Experimentally, we achieved our fastest convergence using ILU(0). In future, we plan to examine the trade-off between more sophisticated preconditioners and total solution time for the forward problem.

2.4 Validation of the Finite Difference Model

In this section, we compare our finite difference diffusion implementation with an integral equation approach using the Rytov approximation, and with a Monte Carlo simulation of photon transport. The experimental setup consists of a slab six cm deep and of various widths. We measure the response due a single source on the top face of the slab using a linear array of 40 detectors, spread over 4 cm. In the examples that follow, we compare normalized magnitude and phase
at the detectors, for a the cases of a homogeneous medium and one containing a spherical absorber.

In the first example, the medium is assumed to be homogeneous, with \( \mu_a = 0.02 \text{ cm}^{-1} \), \( \mu'_s = 10 \text{ cm}^{-1} \), and a modulation frequency of 200 MHz was used. We have kept the relative position of the sources and detectors constant, while varying the width of the intervening slab. The finite difference solution is compared to the analytical solution for a semi-infinite medium, with an refractive-index-matched boundary. The resulting normalized magnitudes and phases at the detectors are shown in Fig. 2.1. Although for the particular case of a slab geometry, the method of images can be used to compute a solution to arbitrary accuracy, this example illustrates the importance of properly modeling the intervening medium.

In order to demonstrate the accuracy of the finite difference solver, we compare the normalized magnitudes obtained using the finite difference method with results obtained by means of a Monte Carlo (MC) simulation comprising \( 10^9 \) photons. In Figures 2.3 and 2.4, the normalized magnitudes and phases at the detectors are shown for the two forward modeling approaches, for the case of a homogeneous cube of width 6 cm. The results clearly show the solutions to be in very good agreement. The finite difference solver, however, is computationally less expensive than MC simulation by a more than two orders of magnitude.

In the next example, we have added an absorbing sphere, 1 cm in diameter, with an absorption coefficient of 0.1 cm\(^{-1}\) between the source and the detectors. In Figures 2.5 and 2.6 we show the magnitude and phase differences, as compared to a homogeneous medium, for the geometries used in the first example. We also compute the expected magnitude and phase differences using the Rylov approximation. The results show that the effect of an absorbing object is highly dependent on the medium geometry, and that it can differ greatly from the Rylov solution.
Figure 2.1: Comparison of the semi-infinite medium analytical solution with the finite difference solution, for slabs of various widths: magnitude
Figure 2.2: Comparison of the semi-infinite medium analytical solution with the finite difference solution, for slabs of various widths: phase
Figure 2.3: Comparison of the MC and FD solutions for a homogeneous cube: magnitude
Figure 2.4: Comparison of the MC and FD solutions for a homogeneous cube: phase
Figure 2.5: Magnitude difference produced by a spherical absorber 1 cm in diameter, with $\mu_a = 0.1\,\text{cm}^{-1}$, for cubes of various widths
Figure 2.6: Magnitude difference produced by a spherical absorber 1 cm in diameter, with $\mu_a = 0.1 \text{ cm}^{-1}$, for cubes of various widths
We also show the magnitude difference caused by a 2.0 cm spherical absorber placed between the source and the detectors, using a modulation frequency of 50 MHz, estimated with the finite difference and Monte Carlo methods. Again we see strong agreement between the simulation approaches, apart from a small discrepancy at the detectors directly opposite the source. We believe that this discrepancy is due to the fact that, strictly speaking, an additional boundary condition needs to be enforced at discontinuities in the optical parameters, specifying the continuity of the component of the photon current normal to the boundary of the discontinuity.
Figure 2.7: Comparison of the magnitude difference, relative to a homogeneous background, caused by the presence of an absorber 2 cm in diameter, with $\mu_a = 0.1 \text{ cm}^{-1}$, computed by means of Monte Carlo and finite difference simulations.
Chapter 3

Nonlinear inversion

3.1 Nonlinear inversion formulation

Although linear inversion is attractive because it can be formulated as the least-squares solution of a large, rank-deficient system, it suffers from a number of disadvantages. First of all, common linearizations, such as the Born and Rytov approximations, are known to be limited in their applicability [14]. In addition, for most realistic geometries, it is not possible to compute the Green functions analytically.

Nonlinear inversion naturally overcomes these difficulties, although we sacrifice being able to easily generalize about existence and uniqueness of solutions. The approach generally taken, and that used in this paper, is to minimize a quadratic function of the mismatch between observed boundary data and boundary data produced by using the forward model with some hypothesized absorption distribution. In statistical terms, we are finding a distribution of absorption (or scattering) that maximizes the likelihood of observing the actual measurements. The inverse problem can be very high-dimensional, resulting in a large computational cost for the optimization. In the section below, we discuss how the Jacobians of the measurements with respect to arbitrarily large parameter sets can be computed efficiently, but even so, solution of the inverse problem requires a large number of solves of the forward problem (with space-varying coefficients), each of
which can be quite time-consuming in itself.

An advantage of using the nonlinear optimization formulation is the ability to arbitrarily transform the measurements into quantities that are more physically meaningful, or that are invariant with respect to various transformations. For example, we will explore the use of normalized measurements, as, for actual DOT experiments, the source power cannot be accurately measured or controlled [50] [51]. As an example of a significant nonlinear transformation, it may be desirable to use magnitude and phase data, reflecting the average intensity and delay at a particular boundary position, repsectively, in the inversion, rather than the real and imaginary parts of the photon fluence, the quantities actually measured.

We assume that for each source we are able to measure the intensity distribution at the detectors, for a total of $NM$ measurements, where we have $N$ sources and $M$ detectors. For source $n$ and detector $m$, the measurement operator used in this paper is a Dirac delta function at the detector position:

$$ h_{n,m} = M_m(\phi_n(r)) = \int_\Omega \phi_n(r) \delta(r - r_m) d^3r $$

(3.1)

In fact it is actually the photon current, or $\hat{n} \cdot \nabla \phi_n(r)$ that is actually measured, but by the partial-current boundary condition, (2.11), the photon density and photon current are assumed to be proportional.

If we can assume multivariate complex Gaussian (or asymptotically Gaussian) noise, the probability of observing the measurement vector, $\mathbf{m}$, given the parameter vectors, $\mathbf{a}$ and $\mathbf{d}$, representing arbitrary absorption and scattering perturbations, is given by the following expression:

$$ P(\mathbf{m}|\mathbf{a}) = \frac{1}{(2\pi)^{D}|\mathbf{R}|} \exp\{-\frac{1}{2}(\mathbf{m} - \mathbf{h(a,d)})^H \mathbf{R}^{-1}(\mathbf{m} - \mathbf{h(a,d)})\} $$

(3.2)

where $D$ is the dimension of the measurement vector, and $\mathbf{R}$ is the covariance matrix of the measurement noise.
Taking the logarithm of (3.2) and ignoring constant terms, we can state the inverse problem as finding absorption and scattering distributions distribution, \( \mathbf{a} \) and \( \mathbf{d} \), that maximize the likelihood of observing \( \mathbf{m} \). Thus, we wish to maximize the following:

\[
l(\mathbf{m}|\mathbf{a}, \mathbf{d}) = -D \log 2\pi - \log |\mathbf{R}| - \frac{1}{2} (\mathbf{m} - \mathbf{h}(\mathbf{a}, \mathbf{d}))^H \mathbf{R}^{-1} (\mathbf{m} - \mathbf{h}(\mathbf{a}, \mathbf{d}))
\]  

(3.3)

where \( \mathbf{m} \) represents a vector of measurements (possibly transformed), and \( \mathbf{h}(\mathbf{a}, \mathbf{d}) \) represents a vector of hypothesized fluence values resulting from \( \mathbf{a} \) and \( \mathbf{d} \), which are infinite-dimensional in the most general case. In the exposition that follows, we will assume that \( \mathbf{R} \) is a diagonal matrix, meaning that measurements by different source-detector pairs are corrupted by independent noise. Here we neglect the unknown source and detector calibration coefficients.

As the general maximum-likelihood problem for an arbitrary absorption distribution is ill-posed, we penalize the likelihood, (3.3) by adding a term that depends on \( \mathbf{a} \) and \( \mathbf{d} \) alone:

\[
l_{\text{pen}}(\mathbf{m}|\mathbf{a}, \mathbf{d}) = -D \log 2\pi - \log |\mathbf{R}| - \frac{1}{2} (\mathbf{m} - \mathbf{h}(\mathbf{a}, \mathbf{d}))^H \mathbf{R}^{-1} (\mathbf{m} - \mathbf{h}(\mathbf{a}, \mathbf{d})) - \lambda S(\mathbf{a}, \mathbf{d})
\]

(3.4)

where \( S(\mathbf{a}, \mathbf{d}) \) is a functional mapping \( \mathbf{a} \) and \( \mathbf{d} \) to \( \mathcal{R} \), often the norm of \( \mathbf{a} \) and \( \mathbf{d} \) or a measure of their average smoothness [17] [23] [12].

In the report that follows we will report results for minimizing the following Weighted Mean-Squared Error (WMSE), which is equivalent to maximizing the likelihood (unpenalized):

\[
WMSE(\mathbf{m}|\mathbf{a}, \mathbf{d}) = (\mathbf{m} - \mathbf{h}(\mathbf{a}, \mathbf{d}))^H \mathbf{R}^{-1} (\mathbf{m} - \mathbf{h}(\mathbf{a}, \mathbf{d}))
\]

(3.5)

In philosophy and in implementation, the penalized-likelihood approach is essentially equivalent to the Bayesian Maximum A-Posteriori (MAP) approach, in which \( \mathbf{a} \) and \( \mathbf{d} \) are taken to be random variables with prior log-distribution
$\lambda S(\mathbf{a}, \mathbf{d})$. Since it is difficult to construct a true probability distribution for arbitrary, infinite-dimensional parameter, an often-used Bayesian imaging technique is to constrain the unknown space-varying absorption to be piecewise constant and to introduce a prior distribution on the changes in value between adjacent regions [26].

In our work, we have implemented the following two smoothing functionals for the absorption [12] [23]:

\[
S_{tk}(\mu_a(\mathbf{r})) = \int_{\Omega} \nabla \mu_a(\mathbf{r}) \cdot \nabla \mu_a(\mathbf{r}) \, d^3 \mathbf{r}
\]  \hspace{1cm} (3.6)

\[
S_{cp}(\mu_a(\mathbf{r})) = \int_{\Omega} \sqrt{\nabla \mu_a(\mathbf{r}) \cdot \nabla \mu_a(\mathbf{r}) + \epsilon} \, d^3 \mathbf{r}
\]  \hspace{1cm} (3.7)

In (3.6), we penalize the average variability of the reconstructed image, as measured by the norm of the gradient, while (3.7) is a smoothing functional that is somewhat more tolerant of sharp edges, $\epsilon$ being a small constant needed in order to ensure differentiability.

### 3.2 Efficient computation of measurement Jacobians

In the imaging approach, the inversion proceeds by estimating the absorption as an arbitrary function of position. In order to accomplish the inversion efficiently, it is necessary to be able to compute the Jacobian of the measurements with respect to a change in absorption at each voxel. While this derivative computation can be easily accomplished for particular discretizations of the forward problem, such as the finite-element and finite-difference methods, it is theoretically satisfying to be able to compute the functional (Fréchet) derivative of the measurements with respect to small perturbations in the absorption.

A functional $J$ (which maps a function $f$ in a normed space to $\mathcal{R}$) is said to be Fréchet differentiable [52] at $f_0$ if it is defined in some sphere $S(f_0)$ about
\[ J(f) = J(f_0) + L(f - f_0) + \|f - f_0\|\zeta(f - f_0) \]  
(3.8)

where \( L \) is a continuous linear (and therefore, bounded) functional, and \( \zeta \) is a function which converges to \( \mathcal{O} \), the zero function, as \( f \to f_0 \).

By the Riesz representation theorem [53], if \( f \) belongs to one of the \( L^p \) function spaces, then the functional \( L(f) \) can be written as \( \int_{\Omega} f g \), where \( g \) is a member of the function space \( L^q \), \( \frac{1}{p} + \frac{1}{q} = 1 \). So, for the case of imaging for the absorption distribution, we wish to find a function \( \nabla l_{\mu_a(r)} \) such that:

\[ l(\mu_a(r) + \delta \mu_a(r)) = l(\mu_a(r)) + \int_{\Omega} \delta \mu_a(r) \nabla l_{\mu_a(r)} \, d^3r \]  
(3.9)

where \( \delta(\mu_a(r)) \) is a small variation in the absorption and \( \zeta(\delta \mu_a(r)) \) in (3.8) is ignored.

The functional derivative of the measurements for an inverse acoustic problem, which is mathematically very similar to that of optical tomography, is derived by Norton [12] using three distinct formulations: the adjoint method, the method of Lagrange multipliers, and the integral equation approach. The adjoint method particularly clearly illustrates how a functional derivative can be found, using variational methods. This approach also introduces an interesting mathematical duality between measurement functions and solutions to the forward problem, using detectors as sources.

Applying the variational approach, we wish to find the variation in the field, \( \delta \phi(r) \) at a particular detector position due to a small variation in the absorption, \( \delta \mu_a(r) \). Specifically, we desire to compute, for source \( n \) and detector \( m \).

\[ \delta h_{n,m} = \int_{\Omega} \delta \phi_n(r) \delta (r - r_m) \, d^3r \]  
(3.10)

where:

\[ -\nabla \cdot D(r) \nabla \phi_n(r) + (\mu_a(r) + \frac{j \omega}{v}) \phi_n(r) = S_n \]  
(3.11)
and
\[-\nabla \cdot D(\mathbf{r}) \nabla (\phi_n(\mathbf{r}) + \delta \phi_n(\mathbf{r})) + (\mu_a(\mathbf{r}) + \delta \mu_a(\mathbf{r}) + \frac{j\omega}{v})(\phi_n(\mathbf{r}) + \delta \phi_n(\mathbf{r})) = S_n \] (3.12)

Subtracting the above equations and discarding second-order terms leads to:
\[(\mu_a(\mathbf{r}) + \frac{j\omega}{v})\delta \phi_n(\mathbf{r}) = \nabla \cdot D(\mathbf{r}) \nabla \delta \phi_n(\mathbf{r}) - \delta \mu_a(\mathbf{r})\phi_n(\mathbf{r}) \] (3.13)

Inserting this expression into (3.10) yields:
\[\delta h_{n,m} = \int_{\Omega} \frac{1}{\mu_a(\mathbf{r}) + \frac{j\omega}{v}}(\nabla \cdot D(\mathbf{r}) \nabla \delta \phi_n(\mathbf{r}) - \delta \mu_a(\mathbf{r})\phi_n(\mathbf{r})) \delta (\mathbf{r} - \mathbf{r}_m) \, d^3 \mathbf{r} \] (3.14)

The key step in the derivation of the adjoint field method is the realization that \(\delta (\mathbf{r} - \mathbf{r}_m)\) in the above equation is equivalent to the diffusion equation with a Robin-type source at the detector location:

\[\delta h_{n,m} = \int_{\Omega} \frac{1}{\mu_a(\mathbf{r}) + \frac{j\omega}{v}}[\nabla \cdot D(\mathbf{r}) \nabla \delta \phi_n(\mathbf{r}) - \delta \mu_a(\mathbf{r})\phi_n(\mathbf{r})] \]
\[\quad \times [-\nabla \cdot D(\mathbf{r}) \nabla \tilde{\phi}_m(\mathbf{r}) + (\mu_a(\mathbf{r}) + \frac{j\omega}{v})\tilde{\phi}_m(\mathbf{r})] \, d^3 \mathbf{r} \] (3.15)

where the \(\tilde{\phi}_m(\mathbf{r})\) is the “adjoint field”:
\[-\nabla \cdot D(\mathbf{r}) \nabla \tilde{\phi}_m(\mathbf{r}) + (\mu_a(\mathbf{r}) + \frac{j\omega}{v})\tilde{\phi}_m(\mathbf{r}) = \delta (\mathbf{r} - \mathbf{r}_m) \] (3.16)

For the particular boundary conditions of the DOT problem, it is shown by Arridge [45], using integration by parts and making a number of additional assumptions, that the functional derivative of a particular frequency-domain source-detector measurement, \(h_{n,m}\), with respect to absorption is the product of \(\phi_n(\mathbf{r})\) for the source and the adjoint field \(\tilde{\phi}_m(\mathbf{r})\) for the detector:

\[\delta h_{n,m} = \int_{\Omega} \delta \mu_a(\mathbf{r})\phi_n(\mathbf{r})\tilde{\phi}_m(\mathbf{r}) \, d^3 \mathbf{r} \] (3.17)
thus:

\[ \nabla_{\mu_a(r)} h_{n,m} = \phi_n(r) \hat{\phi}_m(r) \]  

(3.18)

A similar treatment can be used to derive the functional derivative with respect to scattering perturbations. It can be rigorously shown [54] that the Fréchet derivative computed in the manner shown above is a true derivative, and that \( \phi_n(r) \) is continuous (given appropriate function norms) in \( \delta \mu_a(r) \).

It can be shown using a similar approach [45] that the functional derivative of a measurement with respect to an infinitesimal perturbation in the diffusion coefficient (which we assume to be dependent on the scattering coefficient alone) is

\[ \nabla_{D(r)} h_{nm} = \nabla \phi_n(r) \cdot \nabla \hat{\phi}_m(r) \]  

(3.19)

and the chain rule can be employed to compute the sensitivity of a measurement with respect to a perturbation in scattering.

We note that the equations above yield complex functional gradients, yet we are inverting for real quantities (the absorption and the scattering). We will see in the next section that the Fréchet derivatives of transformed measurements, i.e. of magnitude and phase, are real. In the scattering case, we see that it is the gradient of the perturbation that is significant, which is understandable from an intuitive point of view because the diffusion equation depends on the derivatives of the scattering coefficient, rather than on its actual value.

We see already the tremendous efficiency of the adjoint approach, in that we can compute the the sensitivity of a complex measurement with respect to an arbitrary number of parameters without having to resort to finite differences. We see from (3.18) that the computational and storage costs requirements of a sensitivity calculation are the time of two forward solves, and memory proportional to the number of parameters being estimated.
We now replace the arbitrary function $\mu_a(\mathbf{r})$ by the vector $\mathbf{a}$, ordering our hypothesized measurements as follows:

$$h = [h_{0,0}(\mathbf{a}) \ h_{0,1}(\mathbf{a}) \cdots h_{N-1,M-1}(\mathbf{a})]^T$$

and compute the sensitivity matrix of the data with respect to the parameters for which we are inverting:

$$J_a = \begin{bmatrix}
\nabla_a h_{0,0}(\mathbf{a})^T \\
\nabla_a h_{0,1}(\mathbf{a})^T \\
\cdots \\
\nabla_a h_{N-1,M-1}(\mathbf{a})^T
\end{bmatrix}$$

(3.20)

If we are also inverting for the diffusion coefficient, denoting its discretized space-varying distribution by the vector $d$, we compute the following Jacobian:

$$J_{a,d} = \begin{bmatrix}
\nabla_a h_{0,0}(\mathbf{a})^T & \nabla_a h_{0,0}(d)^T \\
\nabla_a h_{0,1}(\mathbf{a})^T & \nabla_a h_{0,0}(d)^T \\
\cdots & \cdots \\
\nabla_a h_{N-1,M-1}(\mathbf{a})^T & \nabla_a h_{0,0}(d)^T
\end{bmatrix}$$

(3.21)

In general, we do not need to store the entire Jacobian matrix, which would consume storage on the order of $N \times M \times P$, $P$ being the number of parameters we are estimating, a very large quantity in the imaging case. Instead, in optimization algorithms it is generally sufficient to compute the action of the Jacobian, or of its transpose, on a vector. This can be done efficiently by storing the results of $N$ forward solves for the sources, $M$ forward solves for the detectors, and generating rows (or columns) of the Jacobian matrix as needed [51]. In this way, the storage requirement is reduced to $(M + N) \times P$. The storage requirement is the same for inversion for absorption as for inversion for absorption and scattering, as, in both cases, the Jacobian can be computed from the set of forward solutions for the sources and detectors.
It is interesting to note that the Jacobian computed using the adjoint method is identical to the perturbation matrix obtained by means of the Born approximation [14], with each row being essentially a “three-point sensitivity function”. We also note that the Rytov approximation is a particular nonlinear transformation of the measurements that can easily be accommodated within the nonlinear optimization framework. Other transformations, as we will see in the next section, are possible, and can often be useful.

If we now order the measurements in the same way,

\[
m = [m_{0,0} \ m_{0,1} \ \cdots \ m_{N-1,M-1}]^T
\]

the likelihood of the measurements given the discrete set of absorption parameters and asymptotically Gaussian noise can now be expressed as:

\[
l_{\text{pen}}(m|a) = -D \log 2\pi - \log |\Sigma| - \frac{1}{2}(m - h(a))^T \Sigma^{-1} (m - h(a)) - \lambda S_D(a) \tag{3.22}
\]

where \(\Sigma\) is a diagonal matrix representing the measurement noise variances and \(S_D\) is a discrete representation of the smoothing functional (which is, however, still a nonlinear function).

We compute the functional derivative of the smoothing functionals, (3.6) and (3.7) with respect to absorbing perturbations using the same variational approach [12].

\[
\nabla_{\mu_a} (r) S_{tik} = -\nabla^2 \mu_a (r) \tag{3.23}
\]

\[
\nabla_{\mu_a} (r) S_{ep} = \frac{-1}{2\sqrt{\nabla \mu_a (r) \cdot \nabla \mu_a (r)} + \epsilon} \nabla^2 \mu_a (r) \tag{3.24}
\]

Using the above equations and the Jacobian matrix, we can compute the gradient of the likelihood function with respect to the vector of absorption (or scattering) values:

\[
\nabla_a l_{\text{pen}}(m|a) = -\Re \{J_a^H \Sigma^{-1} (m - h(a))\} - \lambda \nabla a S_D(a) \tag{3.25}
\]
We see from the above derivation that the gradient of the penalized likelihood can be computed extremely efficiently. Specifically, computation of the functional gradient requires \(\min(N + M, 2N)\) solutions of the forward problem, as we can compute (3.20) and (3.21) by solving the forward problem for all sources and detectors, with the gradient being a linear combination of \(\phi_n(r)\phi_n(r)\) for all combinations of \(n\) and \(m\). Alternatively, we can compute the Fréchet derivative of the error by solving the forward problem for each source and “backpropagating” [55] the errors at the detectors. In our implementation we use the former approach because, at each iteration of the optimization, the previous iteration’s forward solution can be used as the starting point for the next iteration, somewhat reducing the number of iterations required to achieve the specified tolerance for the forward solution.

The formulation given above can be extended without modification to multiple modulation frequencies by concatenating the observation \((m)\) and model \((h)\) vectors as follows:

\[
[h_{f_1}^T h_{f_2}^T \cdots]^T
\]

(3.26)

\[
[m_{f_1}^T m_{f_2}^T \cdots]^T
\]

(3.27)

where \(f_1, f_2, \ldots\) are the modulation frequencies being used.

The Jacobian matrices can be similarly concatenated, and the results inserted into (3.25) to compute the likelihood of the combined measurements with respect to absorption.

We can also make use of light of varying wavelengths by making use of our knowledge of the extinction coefficients of various chromophores (Oxy-hemoglobin, Deoxy-hemoglobin, water, lipid, etc..) to determine the concentrations of these materials. Below, we assume that there are \(P\) wavelengths being used and that the medium being imaged consists of \(Q\) chromophores. We can compute the
absorption at position \( r \), for wavelength \( \lambda_1 \) by:

\[
\mu^\lambda_1(r) = c_1(r)\varepsilon^\lambda_1 + c_2(r)\varepsilon^\lambda_2 + \ldots + c_P(r)\varepsilon^\lambda_P \tag{3.28}
\]

where \( c_q(r) \) is the concentration of chromophore \( q \) at position \( r \), and \( \varepsilon^\lambda_q \) is the extinction coefficient of chromophore \( q \) for light with wavelength \( \lambda_q \). For multiple wavelengths, we have the following system:

\[
\begin{bmatrix}
\mu^\lambda_1(r) \\
\mu^\lambda_2(r) \\
\vdots \\
\mu^\lambda_P(r)
\end{bmatrix} = \mathbf{E}
\begin{bmatrix}
c_1(r) \\
c_2(r) \\
\vdots \\
c_Q(r)
\end{bmatrix} \tag{3.29}
\]

where \( \mathbf{E} \) is a matrix of extinction coefficients:

\[
\mathbf{E} = 
\begin{bmatrix}
\varepsilon^\lambda_1 & \varepsilon^\lambda_2 & \ldots & \varepsilon^\lambda_Q \\
\varepsilon^\lambda_1 & \varepsilon^\lambda_2 & \ldots & \varepsilon^\lambda_Q \\
\vdots & \vdots & \ddots & \vdots \\
\varepsilon^\lambda_1 & \varepsilon^\lambda_2 & \ldots & \varepsilon^\lambda_Q
\end{bmatrix} \tag{3.30}
\]

In this case, we wish to invert directly for the chromophore concentrations, \( c_1(r), c_2(r), \ldots, c_Q(r) \). As in the case of multiple modulation frequencies, we order the measurements and for wavelengths \( \lambda_1, \lambda_2, \ldots, \lambda_Q \) as follows:

\[
[h^T_{\lambda_1} h^T_{\lambda_2} \ldots h^T_{\lambda_P}]^T \tag{3.31}
\]

\[
[m^T_{\lambda_1} m^T_{\lambda_2} \ldots h^T_{\lambda_P}]^T \tag{3.32}
\]

If we discretize the concentration functions \( c_1(r), c_2(r), \ldots, c_1, c_2, \ldots, \), we can compute the Jacobian of the hypothesized solutions with respect to infinitesimal changes in chromophore concentrations as the following block matrix:

\[
J_e = 
\begin{bmatrix}
J^\lambda_{c_1} & J^\lambda_{c_2} & \ldots & J^\lambda_{c_Q} \\
J^\lambda_{c_1} & J^\lambda_{c_2} & \ldots & J^\lambda_{c_Q} \\
\vdots & \vdots & \ddots & \vdots \\
J^\lambda_{c_1} & J^\lambda_{c_2} & \ldots & J^\lambda_{c_Q}
\end{bmatrix} \tag{3.33}
\]
where $J^\lambda_p$ is the Jacobian matrix for the measurements at wavelength $q$ with respect to the space-varying concentration of chromophore $p$.

From (3.29) we see that this Jacobian can be computed as follows from the Jacobian matrix of each wavelength's measurements with respect to absorption:

$$J_e = (E^T \otimes I_D) \begin{bmatrix} J^\lambda_1 \\ J^\lambda_2 \\ \vdots \\ J^\lambda_D \end{bmatrix}$$

where $D$ is the length of the measurement vector, $\otimes$ is the Kronecker product, and $I_D$ is the identity matrix of size $D$.

### 3.3 Transformation of measured quantities

A significant advantage of the nonlinear reconstruction framework is that transformations of the measured variables can be naturally accommodated. In the case of DOT, the measured quantities are typically the real and imaginary parts of the photon density, acquired at the detectors using a quadrature demodulation scheme. We have implemented two basic classes of transformations - transformations of the measurements into a form that is more physically meaningful than what is actually measured, and transformations that effect various normalizations.

In this section, we denote the multidimensional transformation function by $g$, and if we can make the assumption that the noise in the transformed domain is also Gaussian, our transformed (unpenalized) likelihood function is

$$l(g(m)|a) = -\frac{D}{2} \log 2\pi - \log |\Sigma_g|^\frac{1}{2} - \frac{1}{2}[g(h(a)) - g(m)]^T \Sigma_g^{-1}[g(h(a)) - g(m)]$$

and the functional gradient of this likelihood function is:

$$\nabla_{\mu_0}(r) = -J^T \Sigma_g^{-1}[g(h(a)) - g(m)]$$
where in this case, $J_a$ is the Jacobian of the \textbf{transformed} measurements with respect to the absorption:

$$
J_a = \begin{bmatrix}
\nabla_a g(h(a))_{0,0}^T \\
\nabla_a g(h(a))_{0,1}^T \\
\vdots \\
\nabla_a g(h(a))_{N-1,M-1}^T
\end{bmatrix} \tag{3.37}
$$

In Table 3.3, we list the transformations we have implemented, and the functional derivatives of the transformed measurements are given in Table 3.3. The simplest transformation is to extract the real and imaginary parts of the complex measurements, giving us real Fréchet derivatives. A more complex transformation is to extract the amplitude and phase of each measurement, quantities that are more intuitively meaningful than the real and imaginary measurements. Logarithmic measurements can be of use if we desire all measurements, even those of extremely low amplitude, to be given approximately equal weight. We have introduced the source-normalized error function in response to the observation that it is difficult to measure or control the source power in a DOT system, and that the error functional should therefore be invariant to a scaling of $\sqrt{\sum_{k=0}^{M-1} |h_{ik}|^2}$ and $\sqrt{\sum_{k=0}^{M-1} |m_{ik}|^2}$ for each source $i$.

Once the error gradient has been computed, the nonlinear inversion can be accomplished using standard optimization algorithms. We have implemented the nonlinear conjugate-gradient algorithm [56]:

$$
d^k(r) = -\nabla \mu_a(r) \mathcal{E}^k + \beta_k d^{k-1}(r) \tag{3.38}
$$

$$
\mu_{a}^{k+1}(r) = \mu_{a}^{k}(r) + \alpha_k s^k(r) \tag{3.39}
$$

where $\mathcal{E}^k$ is the error at iteration $k$, $d^k$ is the search direction, $\beta^k$ is a constant computed by Powell’s method, and $\alpha^k$ is a constant such that the mean-squared
<table>
<thead>
<tr>
<th>Transformation</th>
<th>$g_{n,m}(\mathbf{h}(\mu_\alpha(\mathbf{r})))$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Real part</td>
<td>$\frac{1}{2}(h_{n,m} + h_{n,m}^*)$</td>
</tr>
<tr>
<td>Imaginary part</td>
<td>$\frac{1}{2}(h_{n,m} - h_{n,m}^*)$</td>
</tr>
<tr>
<td>Amplitude</td>
<td>$</td>
</tr>
<tr>
<td>Phase</td>
<td>$\tan^{-1}\left(\frac{\Im(h_{n,m})}{\Re(h_{n,m})}\right)$</td>
</tr>
<tr>
<td>Logarithmic Amplitude</td>
<td>$\log</td>
</tr>
<tr>
<td>Source-Normalized Amplitude</td>
<td>$\frac{</td>
</tr>
</tbody>
</table>

Table 3.1: Measurement transformations

<table>
<thead>
<tr>
<th>Transformation</th>
<th>$\nabla_{\mu_\alpha(\mathbf{r})} g_{n,m}(\mathbf{h}(\mu_\alpha(\mathbf{r})))$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Real part</td>
<td>$\Re(\phi_m(\mathbf{r})\tilde{\phi}_n(\mathbf{r}))$</td>
</tr>
<tr>
<td>Imaginary part</td>
<td>$\Im(\phi_m(\mathbf{r})\tilde{\phi}_n(\mathbf{r}))$</td>
</tr>
<tr>
<td>Amplitude</td>
<td>$\frac{1}{</td>
</tr>
<tr>
<td>Phase</td>
<td>$\frac{1}{</td>
</tr>
<tr>
<td>Logarithmic Amplitude</td>
<td>$\frac{1}{</td>
</tr>
<tr>
<td>Source-Normalized Amplitude</td>
<td>$\frac{\sum_{k=0}^{M-1}</td>
</tr>
</tbody>
</table>

Table 3.2: Transformed measurement functional gradients
error decreases from iteration $k - 1$ to iteration $k$. In our implementation, $\alpha^k$ is found by means of a cubic line search.

We impose the constraint that the absorption must be positive by assuming that $\mu_a(\mathbf{r}) = e^{\zeta(\mathbf{r})}$ and estimating $\zeta(\mathbf{r})$ rather than estimating the absorption directly.

### 3.4 Inversion using the adjoint field method - parametric approach

When there is prior information that the medium contains an absorbing anomaly with distinct, sharp boundaries, the imaging approach, in which we estimate the absorption at every point in space within the medium, may be effective. However, it is more natural to invert for the object shape and contrast directly. As discussed by Kilmer et al. [27], this reduced-complexity inversion approach is certainly superior to post-processing an image produced by regularized inversion to extract edges, as standard edge-detection algorithms impose a nonlinear transformation on the estimated image that is divorced from the solution of the inverse problem.

Previous work in this area has included both parametric and non-parametric methods. In this paper we focus on the parametric approach to estimating the shape and location of an absorbing inhomogeneity. As an example of a non-parametric shape-based inversion method, we mention the level set formulation [57], which can model a broad class of object boundaries and has the attractive property of not requiring a priori knowledge of the number of anomalous regions.

A number of parametrizations of an object boundary have been considered in the two-dimensional case, notably Fourier descriptors and B-Splines. However, these parametrizations are not nearly so attractive in the three-dimensional case. For this reason, we explore the example of estimating the location, size, and contrast of an absorbing sphere from frequency-domain optical measurements. As
will be seen in the next section, this problem is particularly difficult for DOT
because perturbations in object size and contrast are almost indistinguishable.
The general inversion formulation and approach in this section extends without
difficulty to more complex parametrizations of the absorber shape.

Thus, we model the absorber as follows:

\[
\mu_a(r) = \mu_a^b + \mu_a^v(r) + \mu_a^o H(O(r))
\]

\[
O(r) = c^2 - (x - x_o)^2 - (y - y_o)^2 - (z - z_o)^2
\]  

(3.40)

where \(\mu_a^b\) is the background absorption, \(\mu_a^v(r)\) is the space-varying component of
the background absorption, \(H\) is the Heaviside step function, \(\mu_a^o\) is the object
absorption contrast, and \(c, x_o, y_o, z_o\) represent the object radius and position.

As in the case of non-parametric inversion, we wish to maximize the likel-
hood of observing the transformed measurements, given a parameter vector \(p\):

\[
l(g(m)|p) = -[g(h(p)) - g(m)]^T \Sigma_g^{-1} [g(h(p)) - g(m)]
\]  

(3.41)

where \(g\) is a measurement transformation, for example one listed in Table 3.3.

In the inversion, we make use of the Jacobian of the measurements with
respect to the object parameters:

\[
J_p = 
\begin{bmatrix}
\frac{\partial g_1(h(p))}{\partial x_o} & \frac{\partial g_1(h(p))}{\partial y_o} & \frac{\partial g_1(h(p))}{\partial z_o} & \frac{\partial g_1(h(p))}{\partial \mu_o^c} & \frac{\partial g_1(h(p))}{\partial c} \\
\frac{\partial g_2(h(p))}{\partial x_o} & \frac{\partial g_2(h(p))}{\partial y_o} & \frac{\partial g_2(h(p))}{\partial z_o} & \frac{\partial g_2(h(p))}{\partial \mu_o^c} & \frac{\partial g_2(h(p))}{\partial c} \\
\frac{\partial g_3(h(p))}{\partial x_o} & \frac{\partial g_3(h(p))}{\partial y_o} & \frac{\partial g_3(h(p))}{\partial z_o} & \frac{\partial g_3(h(p))}{\partial \mu_o^c} & \frac{\partial g_3(h(p))}{\partial c} \\
\vdots & \vdots & \vdots & \vdots & \vdots \\
\end{bmatrix}
\]  

(3.42)
Each term of the above Jacobian can be computed directly, using the adjoint method:

\[
\frac{\partial g_i(h(p))}{\partial c} = \int_{\Omega} 2c \mu_a \delta(O(r)) \nabla_{\mu_a(r)} g_i(h(p)) \; d^3r \\
\frac{\partial g_i(h(p))}{\partial \mu_a} = \int_{\Omega} H(O(r)) \nabla_{\mu_a(r)} g_i(h(p)) \; d^3r \\
\frac{\partial g_i(h(p))}{\partial x_o} = \int_{\Omega} -2(x - x_o) \mu_a \delta(O(r)) \nabla_{\mu_a(r)} g_i(h(p)) \; d^3r \\
\frac{\partial g_i(h(p))}{\partial y_o} = \int_{\Omega} -2(y - y_o) \mu_a \delta(O(r)) \nabla_{\mu_a(r)} g_i(h(p)) \; d^3r \\
\frac{\partial g_i(h(p))}{\partial z_o} = \int_{\Omega} -2(z - z_o) \mu_a \delta(O(r)) \nabla_{\mu_a(r)} g_i(h(p)) \; d^3r
\]

(3.43) (3.44) (3.45) (3.46) (3.47)

where, in the above, \( \delta \) is the three-dimensional Dirac delta function.

In spherical coordinates, with the origin taken to be the center of the absorbing object, we have the following:

\[
\frac{\partial g_i(h(p))}{\partial \mu_a} = \int_{\theta=0}^{2\pi} \int_{\phi=0}^{\pi} \int_{r=0}^{r_c} \nabla_{\mu_a(r)} g_i(h(p)) r^2 \sin \phi \; dr \; d\phi \; d\theta \]

(3.48)

\[
\frac{\partial g_i(h(p))}{\partial c} = 2c^3 \int_{\theta=0}^{2\pi} \int_{\phi=0}^{\pi} \mu_a \nabla_{\mu_a(r)} g_i(h(p)) \sin \phi \; d\phi \; d\theta
\]

(3.49)

\[
\frac{\partial g_i(h(p))}{\partial x_o} = 2c^3 \int_{\theta=0}^{2\pi} \int_{\phi=0}^{\pi} \mu_a \nabla_{\mu_a(r)} g_i(h(p)) \sin^2 \phi \cos \theta \; d\phi \; d\theta
\]

(3.50)

\[
\frac{\partial g_i(h(p))}{\partial y_o} = 2c^3 \int_{\theta=0}^{2\pi} \int_{\phi=0}^{\pi} \mu_a \nabla_{\mu_a(r)} g_i(h(p)) \sin^2 \phi \sin \theta \; d\phi \; d\theta
\]

(3.51)

\[
\frac{\partial g_i(h(p))}{\partial z_o} = 2c^3 \int_{\theta=0}^{2\pi} \int_{\phi=0}^{\pi} \mu_a \nabla_{\mu_a(r)} g_i(h(p)) \cos \phi \sin \phi \; d\phi \; d\theta
\]

(3.52)

Computationally, the adjoint approach to computing the parametric Jacobian is more efficient and accurate than the alternative of using finite differences, particularly since the computational cost is practically independent of the number of parameters.
In order to accomplish the parametric inversion, we use the Levenberg-Marquardt algorithm, which is a variant of the Gauss-Newton approach. This algorithm iteratively estimates the object parameters at iteration \( i + 1 \) using the mismatch between the model and the measurements at iteration \( i \) and the Jacobian of the measurements with respect to the parameters.

\[
\mathbf{p}_{i+1} = \mathbf{p}_i - (J_{p_i}^T \Sigma_g^{-1} J_{p_i} + \lambda I)^{-1} J_{p_i}^T \Sigma_g^{-1} [g(h(p_i)) - g(m)]
\]

If \( \lambda \) is small, the algorithm is essentially the Gauss-Newton method, while, for large \( \lambda \), the algorithm is equivalent to the steepest-descent approach. In our implementation, we vary \( \lambda \) over 14 orders of magnitude, choosing the value that results in the largest reduction of the cost function.

### 3.5 Cramér-Rao Lower Bound

Another advantage of the parametric approach to the inversion problem is that we can explicitly compute a lower bound on the variance of our parameter estimates, which is useful in helping us to determine which problems are solvable, before we attempt to solve them. Assuming that our estimate of the parameter vector, \( \hat{\mathbf{p}} \), is unbiased, a lower bound on its variance is given by [58]:

\[
E[(\hat{\mathbf{p}} - \mathbf{p})(\hat{\mathbf{p}} - \mathbf{p})^T] \geq \mathbf{I_p}
\]

where the information matrix, \( \mathbf{I_p} \) is given by:

\[
\mathbf{I_p} = E_p[(\nabla_p \log f_{g(m)}(g(m)|\mathbf{p})) (\nabla_p \log f_{g(m)}(g(m)|\mathbf{p}))^T]
\]

and the expectation is taken over the distribution of the measurement noise, at the true value of \( \mathbf{p} \). The inequality is meant in the sense that the difference between \( E[(\hat{\mathbf{p}} - \mathbf{p})(\hat{\mathbf{p}} - \mathbf{p})^T] \) and \( \mathbf{I_p} \) is a semi-positive definite matrix.

We have already shown that the gradient of the log-likelihood function with respect to \( \mathbf{p} \) for transformed measurements and Gaussian noise is given by the
following:

\[
\nabla_p \log f_{g(m)}(g(m) | p) = J_p^T \Sigma_g^{-1} (g(m) - g(h(p)))
\]

(3.56)

In the case of a vector of transformed source-detector measurements, the information matrix is given by:

\[
I_p = E_p [J_p^T \Sigma_g^{-1} (g(m) - g(h(p))) (g(m) - g(h(p)))^T \Sigma_g^{-1} J_p]
\]

(3.57)

Taking terms that do not depend on the noise properties out of the expectation gives us:

\[
I_p = J_p^T \Sigma_g^{-1} E_p [(g(m) - g(h(p))) (g(m) - g(h(p)))^T] \Sigma_g^{-1} J_p
\]

(3.58)

Since the quantity that remains in the expectation is just the noise covariance matrix, the final result is:

\[
I_p = J_p^T \Sigma_g^{-1} J_p
\]

(3.59)

We can also compute a bound on variance of the estimate of any particular component of \( p \):

\[
E[(\hat{p}_k - p_k)^2] \geq (I_p^{-1})_{kk}
\]

(3.60)

It is also possible to compute Cramér-Rao bounds on the shape of a parameterized surface [31] [32]. Specifically, if the object boundary is parameterized by \( u \) and \( v \) such that

\[
s(u, v) = s(u, v, \phi) = [x(u, v) \ y(u, v) \ z(u, v)]^T
\]

(3.61)

\[
s(u, v) = B^T(u, v) \phi = \sum_{j=1}^{J} b_j(u, v) \phi_j
\]

(3.62)

where \( b_j(t) \) is the \( j \)’th basis function and \( \phi \) is a parameter vector, then

\[
E[(\hat{s}(u, v) - s(u, v))(\hat{s}(u, v) - s(u, v))^T] \geq B^T(u, v) I_\phi^{-1} B(u, v)
\]

(3.63)
3.6 Noise Modeling

A large proportion of the noise in DOT measurements is a consequence of the fact that a finite, though very large, number of photons is injected into the diffusive medium over the measurement time period. Though the total number of photons is quite enormous, the number that reach any particular detector over a reasonable time period (several minutes, for example) can be very small, for particular source-detector combinations. Also, although the sensitivity functions for a given source $i$ and detector $j$ are the solution of deterministic equations (3.3), they can also be interpreted as the probability that a photon passes through a given volume $\delta V(r')$ on its path from $i$ to $j$ (although proper normalization is necessary) [59]. Likewise, the forward model solutions, although solutions of deterministic partial differential equations, are actually probabilistic in nature.

In this thesis, we will use the shot-noise model for amplitude measurements [38] [60] [51] [61], in which the noise variance is assumed to be proportional to the measurements. Statistically, we are making the assumption that the number of photon arrivals in a given time period is Poisson-distributed, with enough photons arriving that the distribution is asymptotically Gaussian. Specifically, the probability distribution of the magnitude measurement for source $i$ and detector $j$ is:

$$P(|m_{ij}| \mid a) = \frac{1}{\sqrt{2\pi\alpha|h_{ij}|}} \exp \left[ -\frac{(|m_{ij}| - |h_{ij}|)^2}{2\alpha|h_{ij}|} \right]$$  \hspace{1cm} (3.64)

If we assume that the average DC optical power at the detector is:

$$R = \gamma |\phi_{i}^{DC}(r_j)|$$  \hspace{1cm} (3.65)

where $\gamma$ is a constant proportional to the detector area, then

$$\alpha = \frac{2eB}{\kappa \gamma^2}$$  \hspace{1cm} (3.66)
where \( e \) is the charge of an electron, \( B \) is the detector bandwidth, \( \kappa \) is the detector responsivity [60] and \( \beta \) is a coefficient accounting for the fact that the modulation frequency is greater than zero.

The probability distribution, (3.64) is often approximated [51] [60] as:

\[
P(m_{ij} | \mathbf{a}) = \frac{1}{\sqrt{2 \pi \alpha |m_{ij}|}} \exp \left[ - \frac{(|m_{ij}| - |h_{ij}|)^2}{2 \alpha |m_{ij}|} \right]
\]

(3.67)

in order to convert the mean-variance estimation problem into one in which only the mean of \( |h_{ij}| \) is unknown.

A more accurate noise model [61] takes into account the thermal noise in the detector circuitry in addition to the shot noise:

\[
P(m_{ij} | \mathbf{a}) = \frac{1}{\sqrt{2 \pi (\alpha |m_{ij}| + \sigma_T^2)}} \exp \left[ - \frac{(|m_{ij}| - |h_{ij}|)^2}{2 (\alpha |m_{ij}| + \sigma_T^2)} \right]
\]

(3.68)

\[
\sigma_T^2 = \sqrt{kTBF_n}
\]

(3.69)

where \( \sigma_T^2 \) is the thermal noise variance, \( k \) is Boltzman’s constant, \( T \) is the temperature, \( B \) is detector bandwidth, and \( F_n \) is the detector noise level. In order to find the maximum-likelihood solution of the inverse problem, it is not necessary to know \( \alpha \), although it is necessary for statistical testing and computing bounds on parameter estimates.

We will also assume in this thesis that the phase measurements are independent, identically distributed random variables, all with the same variance, following Boas [38].

The above models assume that the source power and detector responsivity (or quantum efficiency), as well as the phase shifts due to source and detector fiber lengths, are known. In actual DOT experiments, however, it is unlikely that these quantities can be known for all source-detector combinations, in part because the actual power transmitted into the medium is dependent on such factors as the placement of fibers, which can vary from experiment to experiment. Boas et.
al. [50] introduced an algorithm to estimate the source and detector coupling coefficients simultaneously with image reconstruction, in the context of a linear model (the Rylov approximation). The extension of this approach to nonlinear models was explored by Oh et. al. [51], of Purdue University who took the source and detector coefficients to be nuisance parameters to be estimated as part of the solution of the inverse problem. Both groups assumed a model in which the source-detector coupling coefficient for measurement $m_{ij}$ is $s_i d_j$, where $s_i$ and $d_j$ are the complex coefficients for source $i$ and detector $j$.

The Purdue group formulated the inverse problem as a Bayesian MAP (Maximum A Prosteriori) estimation problem as follows:

$$\hat{a} = \arg\min_{\alpha, \hat{a}, \hat{s}, \hat{d}} c(\hat{a}, \hat{s}, \hat{d})$$

where $c(\hat{a}, \hat{s}, \hat{d})$ is the cost function to minimize, a function of the absorption estimate, $\hat{a}$ as well as the estimate of the source and detector coefficients:

$$c(\hat{a}, \hat{s}, \hat{d}) = \frac{1}{\alpha}(\mathbf{m} - \text{diag}(\hat{s} \otimes \hat{d}) \mathbf{h}(\hat{a}))^H \Sigma^{-1} (\mathbf{m} - \text{diag}(\hat{s} \otimes \hat{d}) \mathbf{h}(\hat{a})) + D \log \hat{a} + S(\hat{a})$$

(3.71)

In the above, $\otimes$ represents the Kronecker product, $\text{diag}(\mathbf{v})$ is the diagonal matrix such that $(\text{diag}(\mathbf{v}))_i = v_i$, $S(\hat{a})$ is the the prior log-distribution of the absorption estimate $\hat{a}$, $D$ is the dimension of the measurement vector, and $\hat{s}, \hat{d}$ are the vectors of estimated source and detector coefficients. If the shot noise model is used then $\Sigma = \text{diag}([|m_1| |m_2| \ldots |m_D|]^T)$.

We have also developed a doubly stochastic noise model, in which the number of photons that reach detector $j$ from source $i$ over the integration period is assumed to be a Poisson-distributed random variable with mean $s_i^A |\phi_i(r_j)|$, where $|\phi_i(r_j)|$ is the amplitude of the forward model solution for the specified source at the detector location. We also assume that the detector current for each photon arrival is Poisson with parameter $d_j^A$. The net current, $I_{i,j}$, produced is a sum of
a Poisson-distributed number of random variables, each of which is itself Poisson-distributed. It can be shown [62] that the generating function of the aggregate random variable is:

\[ G(z) = e^{s^A |\phi_i(r_j)|} (e^{d_j^A (z-1)} - 1) \]  

(3.72)

and that the variance of the current is:

\[ \text{Var}[I_{i,j}] = s^A_i |\phi_i(r_j)| d_j^A (1 + d_j^A) + \sigma_I^2 \]  

(3.73)

In the case of this noise model, we are optimizing the following cost function:

\[
c(\hat{\alpha}, \vec{\alpha}, \vec{s}^A, \hat{\vec{d}}^A, \vec{d}^A) = \frac{1}{\alpha^A} [g^A(\vec{m}) - \text{diag}(\vec{s}^A \otimes \hat{\vec{d}}^A) g^A(\vec{h}(\vec{\alpha}))]^T (\Sigma_g^A)^{-1} [g^A(\vec{m}) - \text{diag}(\vec{s}^A \otimes \hat{\vec{d}}^A) g^A(\vec{h}(\vec{\alpha}))] + \\
\frac{1}{\alpha^P} [g^P(\vec{m}) + \vec{s}^P + \vec{d}^P - g^P(\vec{h}(\vec{\alpha}))]^T [g^P(\vec{m}) + \vec{s}^P + \vec{d}^P - g^P(\vec{h}(\vec{\alpha}))] + NM \log \hat{\alpha}^P + NM \log \hat{\alpha}^A + S(\vec{\alpha}) \]  

(3.74)

where \( g^A(\vec{v}) \) extracts the amplitude of the vector \( \vec{v} \), \( g^P(\vec{v}) \) extracts its phase, \( \vec{s}^A \) and \( \vec{d}^A \) represent the amplitude coupling coefficients, \( \vec{s}^A \) and \( \vec{d}^A \) are the phase coupling coefficients, for the sources and detectors, respectively, and \( \hat{\vec{s}}^A \) are \( \hat{\vec{d}}^A \) are the amplitude and phase noise levels. The covariance matrix for the amplitude measurements, \( \Sigma_g^A \), is diagonal, with variances given by the modified shot noise model:

\[ \Sigma_g^A = \text{diag}(\vec{s}^A \otimes \vec{f}^A) + \sigma_I^2 I_{NM} \]  

(3.75)

\[ \vec{f}^A = [d_1(d_1 + 1) \ d_2(d_2 + 1) \ \ldots \ d_N(d_N + 1)]^T \]  

(3.76)

We have computed the Jacobian of the cost function with respect to the source and detector coupling coefficients, and estimated all of the coefficients
simultaneously using the Levenberg-Marquardt algorithm. We have also incorporated the estimation of these coefficients into the inversion procedure, using a coordinate descent approach, as described by Oh [51].

3.7 Error Function Analysis

In this section, we use the finite difference forward modeler to analyze, by means of simulations, the cost function, (3.5), for the case of a spherical anomaly in a homogeneous medium. We have computed results for modulation frequencies of 200 MHz and 50 MHz, and background optical properties of $\mu_a = 0.02 \text{ cm}^{-1}$ and $\mu'_s = 5.0 \text{ cm}^{-1}$ were used. The geometry is that of a cube 6 cm on a side, with twenty five sources and twenty five detectors uniformly distributed on the top ($z = 0$ cm) and bottom ($z = 6.0$ cm) faces, respectively. The source-detector geometry and the position of the true absorber is shown in Fig. 3.1. For the simulation experiments in this section, the center of the true absorber is located at $(x = 4.4 \text{ cm}, y = 2.0 \text{ cm}, z = 1.6 \text{ cm})$. The true absorption is $0.2 \text{ cm}^{-1}$ and the object diameter is 2 cm.

For the case of a spherical absorber in a homogeneous medium, an analytical series solution can be found for the scattered photon density [63]:

$$\phi_{sc}(\mathbf{r}_s, \mathbf{r}_d) = \sum_{i=0}^{\infty} A_i h^{(1)}_i(k r_d) Y_i^0(\Omega_d)$$

(3.77)

Here $A_i$ are coefficient determined by applying appropriate boundary conditions at the object boundary, $h^{(1)}_i$ are Hankel functions of the first kind, $Y_i^0(\Omega_d)$ are the spherical harmonics, $k$ is the wavenumber of the homogeneous medium, $\sqrt{\frac{\mu_a + \mu'_s}{D}}$, and $r_d$ is the distance between the sphere boundary and the detector.

It can also be shown that the monopole moment for scattered field due to a spherical absorber in an infinite homogeneous medium is [63]:

$$\phi^{i=0}_{sc} = S \frac{\exp(ikr_s)}{4\pi Dr_s} \frac{\exp(ikr_d)}{4\pi r_d} \left[ \frac{4\pi a^3}{3} \right] \left[ -\frac{\delta \mu_a}{D} \right]$$

(3.78)
Figure 3.1: Source-detector geometry and position of true absorbing sphere
where $S$ is the source intensity, $r_s$ and $r_d$ are the distances from the absorbing sphere's boundary to the source and detector, respectively, $a$ is the sphere radius, and $\delta\mu_a$ is the absorber contrast relative to the background.

Equation 3.78 shows that the monopole moment is proportional to the product of the sphere volume and the absorber contrast. Also, the absorber contrast does not appear in the dipole and quadrupole moments, meaning that higher-order moments are required to give further information about these parameters.

In Figs. 3.2 and ?? we show the change in the magnitude error function as we vary the $x$ and $y$ coordinates of the estimated position of an absorbing sphere. Shot noise with SNR of 40 dB has been added to the measurements. We see from the plot that there is a strong global minimum at the true $(x, y)$ position and that the optimization can be easily accomplished using gradient descent, assuming a reasonable initialization. We have also shown the isocontours of the error function, along with the gradient of the error function with respect to $x_o$ and $y_o$ using (3.46) and (3.47). We see that, as expected, the gradients computed using the adjoint field method are perpendicular to the error isocontours. The corresponding plots for the phase weighted mean-squared error, and an average phase measurement uncertainty of 1°, is shown in Fig. 3.4 and 3.5, for modulation frequencies of 200 MHz and 50 MHz, respectively. We notice that, for the amplitude measurements, there is a strong global minimum at the true object position, while localization with the phase measurements is more difficult, as the cost function is considerably flatter in the vicinity of the true object position.

Figs. 3.6, 3.7, 3.8 and 3.9 shows the result of varying the absorber size and absorption contrast, for the amplitude and phase weighted mean-squared error measures, and modulation frequencies of 200 MHz and 50 MHz, keeping all other parameters fixed. As in the previous example, the amplitude measurements have been corrupted with asymptotically Gaussian shot noise of 40 dB, and the phase
Figure 3.2: The mean-squared magnitude error, for a modulation frequency of 200 MHz, as the estimated \((x, y)\) position of a spherical absorber is varied, with all other parameters assuming their correct values. Also shown are the isocontours of the error function superimposed with vectors in the direction \((\frac{\partial \epsilon}{\partial x_o}, \frac{\partial \epsilon}{\partial y_o})\).
Figure 3.3: The mean-squared magnitude error, for a modulation frequency of 50 MHz, as the estimated \((x, y)\) position of a spherical absorber is varied, with all other parameters assuming their correct values. Also shown are the isocontours of the error function superimposed with vectors in the direction \((\frac{\partial \varphi}{\partial x_0}, \frac{\partial \varphi}{\partial y_0})\).
Figure 3.4: The mean-squared phase error, for a modulation frequency of 200 MHz, as the estimated $(x, y)$ position of a spherical absorber is varied, with all other parameters assuming their correct values. Also shown are the isocontours of the error function superimposed with vectors in the direction $(\frac{\partial \delta}{\partial x_0}, \frac{\partial \delta}{\partial y_0})$. 
Figure 3.5: The mean-squared phase error, for a modulation frequency of 50 MHz, as the estimated $(x, y)$ position of a spherical absorber is varied, with all other parameters assuming their correct values. Also shown are the isocontours of the error function superimposed with vectors in the direction $\left( \frac{\partial \xi}{\partial x_o}, \frac{\partial \xi}{\partial y_o} \right)$.
Figure 3.6: Isocontours of the mean-squared magnitude error, for a modulation frequency of 200 MHz, as the estimated radius and absorption of a spherical absorber are varied, with all other parameters assuming their correct values. Shown along with the direction of \( \left( \frac{\partial E}{\partial \mu}, \frac{\partial E}{\partial c} \right) \) computed using the adjoint-field method.
Figure 3.7: Isocontours of the mean-squared magnitude error, for a modulation frequency of 50 MHz, as the estimated radius and absorption of a spherical absorber are varied, with all other parameters assuming their correct values. Shown along with the direction of \((\frac{\partial E}{\partial \rho}, \frac{\partial E}{\partial c})\) computed using the adjoint-field method.
Figure 3.8: Isocontours of the mean-squared phase error, for a modulation frequency of 200 MHz, as the estimated radius and absorption of a spherical absorber are varied, with all other parameters assuming their correct values. Shown along with the direction of \( \left( \frac{\partial \phi}{\partial r}, \frac{\partial \phi}{\partial \sigma} \right) \) computed using the adjoint-field method.
Figure 3.9: Isocontours of the mean-squared phase error, for a modulation frequency of 50 MHz, as the estimated radius and absorption of a spherical absorber are varied, with all other parameters assuming their correct values. Shown along with the direction of \( \left( \frac{\partial \phi}{\partial \bar{r}}, \frac{\partial \phi}{\partial \bar{c}} \right) \) computed using the adjoint-field method.
uncertainty is $1^\circ$. In contrast to the previous example, we see that the isocontours are approximately hyperbolas, as expected from eq. 3.78. There is in fact a global minimum at the correct parameter values ($\mu_a = 0.3 \text{ cm}^{-1}$, $c = 1.0 \text{ cm}$), but it is very difficult to find using descent-type procedures, as the optimization problem is not convex in these parameters. We see again that the gradients with respect to these two parameters, estimated by means of the adjoint method, are perpendicular to the isocontours.

In Figs. 3.7 we plot the weighted squared error along the hyperbola: $\mu_a r^3 = 0.2$, as a function of radius, for amplitude and phase measurements, respectively, at 200 MHz and 50 MHz. We see that at the given noise levels (40 dB shot noise for the amplitude, and $1^\circ$ phase uncertainty), the phase measurements are better able to distinguish the true global minimum at $r = 1.0$ cm. The phase measurements at along the same hyperbola at 50 MHz, however, are within two standard deviations of the minimum of the cost function, meaning that we have low statistical confidence in the parameter estimates obtained using these measurements.

The conclusion we come to, from examination of the figures is that, for the system geometry and background optical properties in question, at a modulation frequency of 200 MHz, the phase measurements provide useful information both for object localization.

For the given system geometry, background optical properties, and noise variances, we come to the following conclusions:

- At 200 MHz, the phase measurements provide useful information both for object localization and determining object size and contrast.

- At 50 MHz, the phase measurements provide some information for object localization, and almost no information for estimating the object size and
Figure 3.10: WMSE, for along the hyperbola: $\mu_0 r^3 = 0.2$, as a function of object radius. Top left: amplitude measurements at 200 MHz. Top right: phase measurements at 200 MHz. Bottom left: amplitude measurements at 50 MHz. Bottom right: phase measurements at 50 MHz.
contrast.

3.8 Cramér-Rao bound analysis

We have computed the Cramér-Rao bounds for estimating the parameters of a spherical absorber, varying the absorption and radius of the true spherical object over a wide range, for amplitude and phase measurements and the source-detector geometry shown in Fig. 3.1. We assume that the amplitude measurements are corrupted by shot noise with a SNR of 40 dB and the phase measurement uncertainty is 1°. The results for the uncertainty in the estimates of the x and y position and in the depth estimate are shown in Fig. 3.11. As expected from the symmetry of the system, the plots for the uncertainty in the x and y position estimates are very similar to one another, while the depth uncertainty is greater for smaller objects and does not decay as quickly as the true object size increases. Corresponding plots for the uncertainty in the estimates of the object size and contrast are shown in Fig. 3.12. We notice that the uncertainty in the object size estimate is very large for small, low-contrast objects, essentially telling us that this parameter is unresolvable for the given noise level. Qualitatively, these findings agree with those computed by Boas et. al. [64], however we do not have quantitative agreement, most likely due to the fact that we are discretizing our absorbing inclusions, rather than computing their scattered field using the analytical infinite-space solution.

We have also compared the parameter uncertainty error for amplitude measurements only, compared to using both amplitude and phase measurements, at various modulation frequencies. In Fig. 3.13 we compare the radius estimation error for amplitude measurements alone with measurements of both amplitude and phase, as a function of the object absorption, where the modulation frequency is 0 MHz, 50 MHz, and 200 MHz. The true object radius is 1 cm. We see that
using both the amplitude and phase data allows for much better estimation of the object size for the given parameters than using the amplitude measurements alone, and that the estimation uncertainty is smaller for 200 MHz measurements of the amplitude and phase, as opposed to measurements at 50 MHz. This result is somewhat contrary to the results presented by Boas [64], where it is stated that the optimal frequency for characterizing absorbing objects is 0 MHz. However, we have used a different source-detector geometry and a considerably greater number of measurements in order to obtain our result.
Figure 3.11: The Cramér-Rao lower bound for estimating the x (left) and y (right) position and depth (bottom) of an absorbing sphere, as we vary the sphere radius and absorption
Figure 3.12: The Cramér-Rao lower bound for estimating the radius (left) and absorption (right) of an absorbing sphere, as we vary the sphere radius and absorption
Figure 3.13: Uncertainty in the radius estimate, where the true radius is 1 cm, as a function of object absorption, comparing amplitude measurements only with amplitude and phase measurements, for a number of modulation frequencies.
Chapter 4

Inversion results

4.1 Simulation results

In this section, we show simulation results for non-parametric and parametric inversion where the true object is a an absorbing sphere with varying values of $\mu_a$ centered at $(x = 4.4 \text{ cm}, \ y = 2.0 \text{ cm}, \ z = 1.6 \text{ cm})$ with a diameter of 2 cm. We repeat the experiment with varying levels of measurement amplitude and phase noise, demonstrating the sensitivity of the parameters to measurement uncertainty. The background absorption is $0.02 \text{ cm}^{-1}$ and the background $\mu'_a$ is $5.0 \text{ cm}^{-1}$. A transmission geometry is used for the source-detector measurements, and the modulation frequency is 200 MHz. The location of the sources and detectors is shown in Fig. 4.1. The sources are located in the $z = 0.2 \text{ cm}$ plane, and the detectors are located on the $z = 6.0 \text{ cm}$ plane.

4.1.1 Nonparametric Inversion

Here we show the reconstruction of an absorbing sphere, making use of amplitude measurements only, with a shot noise level of 40 dB. An edge-preserving regularizer was used.

We show various slices of the of the true solution in Fig. 4.2. In Fig. 4.3 we show corresponding slices, where the absorption of the actual object is $0.3 \text{ cm}^{-1}$. We have the repeated the simulation for the case where the true object absorption
Figure 4.1: Source-detector geometry and absorber location for inversion experiments
Figure 4.2: True solution for an absorbing sphere with $\mu_a = 0.3 \text{cm}^{-1}$ in a homogeneous background.
Figure 4.3: Reconstructed non-parametric solution for an absorbing sphere with $\mu_a = 0.3\text{cm}^{-1}$ using amplitude measurements
Figure 4.4: Reconstructed non-parametric solution for an absorbing sphere with \( \mu_a = 0.2 \text{ cm}^{-1} \), using amplitude measurements
<table>
<thead>
<tr>
<th>Noise Level</th>
<th>Final WMSE</th>
<th>Penalty</th>
</tr>
</thead>
<tbody>
<tr>
<td>20 dB</td>
<td>835.37</td>
<td>43.35</td>
</tr>
<tr>
<td>40 dB</td>
<td>4446.71</td>
<td>32.55</td>
</tr>
</tbody>
</table>

Table 4.1: Final WMSE and penalty when true object $\mu_a = 0.3 \text{ cm}^{-1}$

is 0.2 cm$^{-1}$, and show slices of the reconstruction in Fig. 4.4. The reconstructions look very similar, localizing the object very well, but it is difficult to quantify the true object absorption using this approach. We also note the image artifacts in the $z = 1$ cm plane, near the sources.

Tables 4.1.1 and 4.1.1 show the final negative log-likelihood and penalty terms for inversion of absorbing spheres with $\mu_a = 0.3 \text{ cm}^{-1}$ and $\mu_a = 0.2 \text{ cm}^{-1}$, respectively. We note that the final weighted mean-squared errors (WMSE) for the the higher SNR data are quite far from their expected minimum value of 625 (the number of measurements).

<table>
<thead>
<tr>
<th>Noise Level</th>
<th>Final WMSE</th>
<th>Penalty</th>
</tr>
</thead>
<tbody>
<tr>
<td>20 dB</td>
<td>725.49</td>
<td>37.38</td>
</tr>
<tr>
<td>40 dB</td>
<td>2122.05</td>
<td>50.61</td>
</tr>
</tbody>
</table>

Table 4.2: Final WMSE and penalty when true object $\mu_a = 0.2 \text{ cm}^{-1}$
4.1.2 Parametric Inversion

We repeated the simulation experiment in the previous example, placing the center of an absorbing sphere with various levels of absorption at the position \((x = 4.4 \text{ cm}, y = 2.0 \text{ cm}, z = 1.6 \text{ cm})\) in a material with background absorption of 0.02 cm\(^{-1}\) and background reduced scattering coefficient of 5 cm\(^{-1}\).

In the inversion, we assume that we have prior information that the true object is a sphere, and we use the parametric inversion approach described above to estimate its parameters. For the initial guess, we use a sphere with \(\mu_a = 0.2 \text{ cm}^{-1}\) centered at \((x = 3.0 \text{ cm}, y = 3.0 \text{ cm}, z = 3.0 \text{ cm})\).

We examined the use of a number of optimization algorithms for the parametric estimation. We compare Levenberg-Marquardt (LM) nonlinear regression, Modified Levenberg-Marquardt (MLM) inversion, and Modified Coordinate Descent (MCD). The LM algorithm optimizes for all parameters simultaneously, with no accommodation made for the fact that the cost function is not convex with respect to several of the parameters. The MLM algorithm attempts to deal with this difficulty by periodically (every 10 iterations) conducting a hyperbolic line search along the curve \(\mu_a \hat{r}^3 = c\), \(c\) being a constant. The MCD algorithm conducts Levenberg-Marquardt inversion for the positional parameters separately from the object radius and absorption, alternating between the two sets of parameters and also periodically performing a hyperbolic line search. For this comparison, the true object absorption was 0.3 cm\(^{-1}\). In the inversion, we made use of amplitude measurements, with asymptotically Gaussian shot noise at a level of 40 dB. The progression of the algorithm is shown in Fig. 4.5 We notice that all of the optimization approaches minimize the weighted mean-squared error and the positional error fairly rapidly, although convergence for the coordinate descent is somewhat slower than for the algorithms that update all parameters at once. We
also notice that the LM algorithm gives object absorption and radius estimates that are relatively far from their true values, compared to the MLM and MCD approaches. Table 4.1.2 gives the final WMSE and object parameter estimates for the three algorithms. In terms of WMSE, however, the inversion results can be considered to be identical, since, in all cases, the final WMSE is within two standard deviations of its expected value of 625, as the WMSE is a central $\chi^2_{625}$ random variable which is asymptotically distributed as $\mathcal{N}(625, 625)$. 
<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Final WMSE</th>
<th>$x$ (4.4)</th>
<th>$y$ (2.0)</th>
<th>$z$ (1.6)</th>
<th>$\mu_a$ (0.3)</th>
<th>$r$ (1.0)</th>
</tr>
</thead>
<tbody>
<tr>
<td>LM</td>
<td>609.78</td>
<td>4.49</td>
<td>1.98</td>
<td>1.51</td>
<td>0.14</td>
<td>1.23</td>
</tr>
<tr>
<td>MLM</td>
<td>645.155</td>
<td>4.44</td>
<td>1.99</td>
<td>1.59</td>
<td>0.21</td>
<td>1.09</td>
</tr>
<tr>
<td>MCD</td>
<td>611.174</td>
<td>4.42</td>
<td>1.99</td>
<td>1.52</td>
<td>0.20</td>
<td>1.10</td>
</tr>
</tbody>
</table>

Table 4.3: Final Weighted Mean-Squared Error and parameter estimates for various algorithms

We have also examined the effect of an object’s absorption on the ability of the nonlinear regression approach to characterize its position, absorption, and shape, for various levels of additive noise. As in the previous examples, the true object is a sphere centered at ($x = 4.4$ cm, $y = 2.0$ cm, $z = 1.6$ cm), in a homogeneous background. These idealized examples are useful for determining the absolute limits of object detection and characterization. We use amplitude measurements only and show the convergence of the object WMSE, position, sphere radius, and absorption for additive Gaussian shot noise levels of 80 dB, 60 dB, 40 dB, and 20 dB. The results for an absorbing sphere with $\mu_a = 0.3$ cm$^{-1}$ are shown in Fig. 4.6 For an object with $\mu_a = 0.2$ cm$^{-1}$, the results are shown in Fig. 4.7 and the results for an object with $\mu_a = 0.1$ are shown in Fig. 4.8. The quantitative results are given in Tables 4.1.2, 4.1.2, and 4.1.2. We notice that for the extremely low-noise case (80 dB SNR), the algorithm tends to stop at a local minimum, not finding the true global minimum of the cost function, which is $625 \pm 3(25)$, as the standard deviation of a $\chi^2_{625}$ random variable is approximately 25. We also notice that the object localization is extremely accurate in all cases, achieving essentially millimeter accuracy. The situation is quite different, however, for the estimation of the object radius and absorption. This is only accurately accom-
Figure 4.5: Parametric errors, as a function of iteration, for several optimization approaches, using amplitude measurements.
plished consistently for the object with absorption $0.3 \text{ cm}^{-1}$. We also see that, since the cost of the estimated solution is, in all cases, within approximately two standard deviations of the known minimum, it may not be possible to estimate these parameters accurately, using amplitude data only.

We also repeated the above parametric inversion simulation experiment, for an absorber with $\mu_a = 0.2 \text{ cm}^{-1}$, now making use of the phase measurements, and varying the signal-to-noise ratio of the measurements. The results are shown in Fig. 4.9, and the quantitative results are given in Table 4.1.2. The weighted error values are higher than those in Tables 4.1.2, 4.1.2, and 4.1.2 because we have doubled the number of measurements, with the mean value of the cost function now being 1250. We see that making use of the phase measurements significantly improves our ability to simultaneously estimate the sphere size and absorption for the geometry in question.

<table>
<thead>
<tr>
<th>Noise Level</th>
<th>Final WMSE</th>
<th>$x$ (4.4)</th>
<th>$y$ (2.0)</th>
<th>$z$ (1.6)</th>
<th>$\mu_a$ (0.3)</th>
<th>$r$ (1.0)</th>
</tr>
</thead>
<tbody>
<tr>
<td>80 dB</td>
<td>786.4</td>
<td>4.44</td>
<td>1.98</td>
<td>1.55</td>
<td>0.20</td>
<td>1.1</td>
</tr>
<tr>
<td>60 dB</td>
<td>634.63</td>
<td>4.40</td>
<td>2.0</td>
<td>1.62</td>
<td>0.29</td>
<td>1.0</td>
</tr>
<tr>
<td>40 dB</td>
<td>626.47</td>
<td>4.42</td>
<td>1.99</td>
<td>1.61</td>
<td>0.28</td>
<td>1.0</td>
</tr>
<tr>
<td>20 dB</td>
<td>595.4</td>
<td>4.41</td>
<td>2.01</td>
<td>1.52</td>
<td>0.42</td>
<td>0.88</td>
</tr>
</tbody>
</table>

Table 4.4: Final Weighted Mean-Squared Error and parameter estimates for an absorbing sphere with $\mu_a = 0.3 \text{ cm}^{-1}$ for amplitude measurements.
Figure 4.6: Parametric errors when the true object $\mu_a = 0.3 \text{ cm}^{-1}$, using amplitude measurements
Figure 4.7: Parametric errors when the true object $\mu_\alpha = 0.2 \text{ cm}^{-1}$, using amplitude measurements
Figure 4.8: Parametric errors when the true object $\mu_a = 0.1 \text{ cm}^{-1}$, using amplitude measurements
<table>
<thead>
<tr>
<th>Noise Level</th>
<th>Final WMSE</th>
<th>$x$ (4.4)</th>
<th>$y$ (2.0)</th>
<th>$z$ (1.6)</th>
<th>$\mu_a$ (0.2)</th>
<th>$r$ (1.0)</th>
</tr>
</thead>
<tbody>
<tr>
<td>80 dB</td>
<td>904.06</td>
<td>4.38</td>
<td>2.0</td>
<td>1.54</td>
<td>0.18</td>
<td>1.0</td>
</tr>
<tr>
<td>60 dB</td>
<td>654.07</td>
<td>4.34</td>
<td>2.03</td>
<td>1.65</td>
<td>0.65</td>
<td>0.71</td>
</tr>
<tr>
<td>40 dB</td>
<td>626.47</td>
<td>4.42</td>
<td>1.99</td>
<td>1.61</td>
<td>0.28</td>
<td>1.0</td>
</tr>
<tr>
<td>20 dB</td>
<td>570.91</td>
<td>4.37</td>
<td>2.04</td>
<td>1.64</td>
<td>0.64</td>
<td>1.02</td>
</tr>
</tbody>
</table>

Table 4.5: Final Weighted Mean-Squared Error and parameter estimates for an absorbing sphere with $\mu_a = 0.2\,\text{cm}^{-1}$, using amplitude measurements

<table>
<thead>
<tr>
<th>Noise Level</th>
<th>Final WMSE</th>
<th>$x$ (4.4)</th>
<th>$y$ (2.0)</th>
<th>$z$ (1.6)</th>
<th>$\mu_a$ (0.1)</th>
<th>$r$ (1.0)</th>
</tr>
</thead>
<tbody>
<tr>
<td>80 dB</td>
<td>791.63</td>
<td>4.33</td>
<td>2.03</td>
<td>1.65</td>
<td>0.28</td>
<td>0.7</td>
</tr>
<tr>
<td>60 dB</td>
<td>641.02</td>
<td>4.35</td>
<td>2.03</td>
<td>1.64</td>
<td>0.19</td>
<td>0.8</td>
</tr>
<tr>
<td>40 dB</td>
<td>678.84</td>
<td>4.44</td>
<td>1.99</td>
<td>1.61</td>
<td>0.39</td>
<td>0.64</td>
</tr>
<tr>
<td>20 dB</td>
<td>571.43</td>
<td>4.46</td>
<td>2.04</td>
<td>1.76</td>
<td>0.56</td>
<td>0.62</td>
</tr>
</tbody>
</table>

Table 4.6: Final Weighted Mean-Squared Error and parameter estimates for an absorbing sphere with $\mu_a = 0.1\,\text{cm}^{-1}$, using amplitude measurements
Figure 4.9: Parametric errors when the true object $\mu_a = 0.2 \text{ cm}^{-1}$, amplitude and phase measurements
<table>
<thead>
<tr>
<th>Noise Level</th>
<th>Final WMSE</th>
<th>$x$ (4.4)</th>
<th>$y$ (2.0)</th>
<th>$z$ (1.6)</th>
<th>$\mu_a$ (0.2)</th>
<th>$r$ (1.0)</th>
</tr>
</thead>
<tbody>
<tr>
<td>80 dB</td>
<td>1262.68</td>
<td>4.40</td>
<td>2.01</td>
<td>1.61</td>
<td>0.20</td>
<td>0.99</td>
</tr>
<tr>
<td>60 dB</td>
<td>1220.01</td>
<td>4.37</td>
<td>2.02</td>
<td>1.63</td>
<td>0.34</td>
<td>0.83</td>
</tr>
<tr>
<td>40 dB</td>
<td>1249.38</td>
<td>4.43</td>
<td>1.99</td>
<td>1.63</td>
<td>0.15</td>
<td>1.06</td>
</tr>
<tr>
<td>20 dB</td>
<td>1167.41</td>
<td>4.43</td>
<td>1.96</td>
<td>1.58</td>
<td>0.16</td>
<td>1.04</td>
</tr>
</tbody>
</table>

Table 4.7: Final Weighted Mean-Squared Error and parameter estimates for an absorbing sphere with $\mu_a = 0.2\text{cm}^{-1}$, amplitude and phase measurements
4.2 Parametric inversion in the case of model mismatch

While it would be convenient for our algorithm if all anomalies were spherical in shape, it is clearly not often that they actually are. To examine our algorithm's robustness and convergence in the case of model mismatch, we have conducted an experimental study in which we attempt to reconstruct ellipsoids of varying eccentricities. The background optical parameters, modulation frequency, and source-detector geometry are the same as in the previous section. The true objects are centered at \( (x = 4.4 \text{ cm}, y = 2.0 \text{ cm}, z = 1.6 \text{ cm}) \), and have \( \mu_a = 0.3 \text{ cm}^{-1} \).

We vary the length of the major axis of the objects (the y axis) from 2.4 cm to 3.6 cm, while the length of the remaining axes is fixed at 1.0 cm. The true objects are shown in Fig. 4.10. The true and estimated absorptions for an imaging reconstruction of an ellipsoid with major axis length 2.4 cm are shown in Fig. 4.11, and the results in the case where the true absorber is an ellipsoid with major axis length 3.6 cm are shown in Fig. 4.12. Numerical results for the imaging-type reconstructions are shown in Table 4.2. We have also run simulations in which we attempt to find the best-fit sphere for each of the ellipsoids, using amplitude measurements, with 40 dB of Gaussian shot noise added to the measurements. The results are given in Table 4.2. We notice that the WMSE of the best parameter estimate increases significantly as the object becomes less spherical, giving us a criterion for deciding that when a more sophisticated parametric model would be appropriate. We also notice that the radius of the best-fit sphere increases as the true object size increases, while we are able to estimate the true absorption accurately in all cases.
Figure 4.10: Source-detector geometry and positions of ellipsoids with major axis lengths of 2.4 cm, 2.8 cm, 3.2 cm, and 3.6 cm, clockwise, from top left
Figure 4.11: True and estimated absorption for ellipsoid with major axis length 2.4 cm

Figure 4.12: True and estimated absorption for ellipsoid with major axis length 3.6 cm
<table>
<thead>
<tr>
<th>Noise Level</th>
<th>Final WMSE</th>
<th>Penalty</th>
</tr>
</thead>
<tbody>
<tr>
<td>Major axis 2.4 cm</td>
<td>42668.6</td>
<td>51.59</td>
</tr>
<tr>
<td>Major axis 2.8 cm</td>
<td>28298.3</td>
<td>71.24</td>
</tr>
<tr>
<td>Major axis 3.2 cm</td>
<td>114598.0</td>
<td>49.36</td>
</tr>
<tr>
<td>Major axis 3.6 cm</td>
<td>88759.7</td>
<td>62.00</td>
</tr>
</tbody>
</table>

Table 4.8: Final WMSE and penalty for imaging-type absorption reconstructions when the true objects are absorbing ellipsoids of various eccentricities.

<table>
<thead>
<tr>
<th>Noise Level</th>
<th>Final WMSE</th>
<th>$x$ (4.4)</th>
<th>$y$ (2.0)</th>
<th>$z$ (1.6)</th>
<th>$\mu_a$ (0.3)</th>
<th>$r$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Major axis 2.4 cm</td>
<td>755.27</td>
<td>4.21</td>
<td>1.8</td>
<td>1.47</td>
<td>0.29</td>
<td>1.07</td>
</tr>
<tr>
<td>Major axis 2.8 cm</td>
<td>1159.7</td>
<td>4.20</td>
<td>1.84</td>
<td>1.51</td>
<td>0.31</td>
<td>1.10</td>
</tr>
<tr>
<td>Major axis 3.2 cm</td>
<td>2244.57</td>
<td>4.23</td>
<td>1.89</td>
<td>1.63</td>
<td>0.28</td>
<td>1.18</td>
</tr>
<tr>
<td>Major axis 3.6 cm</td>
<td>3879.17</td>
<td>4.21</td>
<td>1.96</td>
<td>1.64</td>
<td>0.30</td>
<td>1.20</td>
</tr>
</tbody>
</table>

Table 4.9: Final WMSE and parameter estimates where the true objects are absorbing ellipsoids of various eccentricities.
4.3 Phantom Results

We demonstrate the reconstruction using phantom data collected by Quan Zhang and colleagues at the DOT lab of the Massachusetts General Hospital, using their frequency-domain instrument. The experiment consists of a homogeneous cube with absorption $0.023 \text{ cm}^{-1}$ and $\mu'_a = 10 \text{ cm}^{-1}$ containing a spherical chamber mixed with india ink to achieve an absorption of $\mu_a = 0.3 \text{ cm}^{-1}$. The object is probed with 780 nm light injected at 30 sources and measured at nine detectors. A modulation frequency of 70 MHz was used. The experimental setup and the configuration of the sources and detectors is shown in fig. 4.13. A reference frame was collected, without an absorber present, in order to determine the source-detector coupling coefficients.

The nonlinear absorption reconstruction, using logarithmic amplitude data, is shown in fig. 4.14. The source-detector coupling coefficients were obtained from a baseline measurement in a homogeneous medium. The reconstruction localizes the absorber fairly well, but there are image artifacts, and the actual absorption of the inhomogeneity differs from the values obtained by the reconstruction. We believe the reason for this discrepancy is that, as seen in the previous section, it is difficult to distinguish a larger, less absorbing, object from a smaller, more absorbing, object. We have also computed a reconstruction from the same data, in which the source-detector coupling coefficients were not known a priori, but were rather estimated as part of the nonlinear inversion process. The results, for amplitude measurements and a doubly stochastic noise model, are shown in Fig. 4.15.

With the same data we also conducted a parametric reconstruction, again using logarithmic amplitude data, using our prior knowledge that the true absorber is spherical. The decrease in the WMSE is shown in figure 4.16. We show
Figure 4.13: Experimental setup for phantom consisting of a spherical absorber in a homogeneous background
Figure 4.14: Nonlinear absorption reconstruction of phantom using log-magnitude data
Figure 4.15: Nonlinear absorption reconstruction of phantom, with unknown source-detector weights estimated iteratively, using amplitude measurements and a doubly stochastic noise model.
the positional error and the error in estimating the object radius in fig. 4.17. It is seen that the algorithm rapidly converges to the correct parameter values for the object size and position. The absorption, however, is considerably over-estimated, with an estimated value of $\mu_a = 0.89 \text{ cm}^{-1}$. 
Figure 4.16: MSE decrease as a function of iteration using parametric reconstruction
Figure 4.17: Error in estimating the position and radius of an absorbing sphere as a function of iteration.
Chapter 5

Conclusion and Future Work

We have presented a nonlinear optimization framework for Diffuse Optical Tomography, a technology with increasing promise for a wide variety of medical applications. The essential problem is to image spatial variations in optical absorption and scattering, possibly at multiple wavelengths, in a highly scattering medium, using frequency-domain measurements of modulated light. These variations can be used to infer physiologically significant information, such as tissue oxygen saturation and tissue density. We assume that we have data from a relatively small number of discrete source and detector positions on the surface of the medium, and we wish to reconstruct the absorption and scattering within a three-dimensional volume.

In the frequency-domain, the inverse problem is essentially a near-field imaging problem, in which the wavelength of the Diffuse Photon Density Waves is often an order of magnitude or more larger than the features being imaged. In addition, we do not assume that perturbations in absorption and scattering are small, and thus the problem is nonlinear in the unknown parameters.

We have developed finite-difference frequency-domain software for the rapid solution of the photon density within a medium given known optical parameters. We have also shown that functional Jacobians of the boundary measurements can be very efficiently computed using the adjoint method. We have presented a
number of transformations that can be applied to the boundary measurements, for example, in order that the measurements be invariant to source power. We have also given expressions for the Jacobians of the transformed measurements. We have shown how various regularizers, notably the edge-preserving regularizer, can be incorporated in the nonlinear inversion framework. The same approach was also used in the case where the object shape function is known. We computed the gradients of the size, contrast and position parameters where the actual inhomogeneity is an absorbing sphere, again using the adjoint approach. Analysis of the cost function with respect to estimates of the object parameters showed that the estimation problem is not convex in the object size and contrast. In addition, we computed statistical lower bounds on our ability to estimate the object parameters, showing that, in general, estimation accuracy increases as the object size and absorption contrast increase. Our data also suggest that measurements of amplitude and phase can give superior performance to amplitude measurements alone, with the object characterization uncertainty dependent on modulation frequency.

We see a number of possible directions for future research. First, the parametric approach generalizes to more complex parametrizations of the object boundary. It would be interesting to investigate regularization approaches for more complex object functions, for example, algorithms for optimally choosing the number of control points for three-dimensional splines, or the number of terms to use in a spherical harmonic expansion. It would also be of interest to investigate statistical bounds on our ability to estimate an object’s shape, for these more complex parametrizations, and for various measurement geometries. It would also be of use to analyze the influence of an inhomogeneous background on shape estimation accuracy.
Bibliography


[57] “A level-set approach for inverse problems involving level sets,” .


