IDENTIFYING AND ADDRESSING THE ERROR SOURCES IN DIFFUSE OPTICAL TOMOGRAPHY

By

Murat Güven

A Thesis Submitted to the Graduate
Faculty of Rensselaer Polytechnic Institute
in Partial Fulfillment of the
Requirements for the Degree of
DOCTOR OF PHILOSOPHY
Major Subject: Electrical, Computer, and Systems Engineering

Approved by the
Examining Committee:

Birsen Yazıcı, Thesis Adviser

Eldar Giladi, Member

Badrinath Roysam, Member

Qiang Ji, Member

Xavier Intes, Member

Rensselaer Polytechnic Institute
Troy, New York

December 2007
(For Graduation December 2007)

IDENTIFYING AND ADDRESSING THE ERROR SOURCES IN DIFFUSE OPTICAL TOMOGRAPHY

By

Murat Güven

An Abstract of a Thesis Submitted to the Graduate

Faculty of Rensselaer Polytechnic Institute

in Partial Fulfillment of the

Requirements for the Degree of

DOCTOR OF PHILOSOPHY

Major Subject: Electrical, Computer, and Systems Engineering

The original of the complete thesis is on file
in the Rensselaer Polytechnic Institute Library

Examining Committee:

Birsen Yazıcı, Thesis Adviser
Eldar Giladi, Member
Badrinath Roysam, Member
Qiang Ji, Member
Xavier Intes, Member

Rensselaer Polytechnic Institute
Troy, New York

December 2007
(For Graduation December 2007)

CONTENTS

LIST OF TABLES vii

LIST OF FIGURES ix

ACKNOWLEDGMENT xi

ABSTRACT xiii

1. Introduction 1
    - 1.1 DOT in a Clinical Perspective 2
    - 1.2 Diffuse Optical Tomography as an Imaging Problem 3
    - 1.3 Thesis Outline 9
        - 1.3.1 Effect of discretization on the accuracy of optical absorption imaging 10
        - 1.3.2 Adaptive mesh generation for DOT 12
        - 1.3.3 Effect of linearization by Born approximation on the accuracy of DOT imaging 14
        - 1.3.4 Effect of discretization on the simultaneous reconstruction of optical absorption and diffusion coefficients 15

2. Effect of discretization error in diffuse optical absorption imaging 19
    - 2.1 Introduction 19
    - 2.2 Forward and Inverse Problems 21
        - 2.2.1 Forward Problem 22
        - 2.2.2 Inverse Problem 25
    - 2.3 Discretization of the Inverse and Forward Problems 27
        - 2.3.1 Inverse Problem Discretization 27
        - 2.3.2 Forward Problem Discretization 28
        - 2.3.3 Discretization of the Inverse Problem with Operator Approximations 29
    - 2.4 Discretization-based Error Analysis 29
        - 2.4.1 Case 1: The kernel $\kappa(\mathbf{x}, \acute{\mathbf{x}})$ is exact 30
        - 2.4.2 Case 2: The kernel is degenerate 35
        - 2.4.3 Iterative Born Approximation 42
    - 2.5 Conclusion 43

3. Adaptive mesh generation in diffuse optical absorption imaging 45

3.1 Introduction 45

3.2 Overview 48

3.2.1 Forward and inverse problems in DOT 48

3.2.2 Effect of inverse problem discretization 50

3.2.3 Effect of forward problem discretization 51

3.3 Adaptive Mesh Generation 53

3.3.1 Adaptive Mesh Generation for the Forward Problem 53

3.3.2 Computational cost of the adaptive mesh generation algorithm for the forward problem 55

3.3.3 Adaptive mesh generation for the inverse problem: 56

3.3.4 Computational cost of the adaptive mesh generation algorithm for the inverse problem 58

3.4 Numerical Experiments 59

3.4.1 Simulation Study 1 60

3.4.2 Simulation Study 2 63

3.4.3 Simulation Study 3 65

3.5 Conclusion 70

4. Error in Optical Absorption Images due to Born Approximation in Diffuse Optical Tomography 78

4.1 Introduction 78

4.2 The Scattered Field and Born Approximation 80

4.3 Inverse Problem 82

4.3.1 Inverse problem statement without Born approximation 83

4.3.2 Inverse problem based on Born approximation 84

4.4 Analysis of the error resulting from Born approximation and Tikhonov regularization 86

4.4.1 Effect of Born approximation 86

4.4.2 Effect of Tikhonov regularization 91

4.4.3 The overall error 93

4.5 Discussion 93

4.6 Numerical Experiments 94

4.6.1 Simulation 1 96

4.6.2 Simulation 2 98

4.7 Conclusions 99

5. Effect of discretization on the accuracy of simultaneously reconstructed absorption and scattering images 107

5.1 Forward and Inverse Problems 109

5.1.1 Forward Problem 110

5.1.2 Inverse Problem 112

5.1.3 Regularization of the inverse problem 113

5.1.4 Existence and boundedness of the inverse operator 115

5.2 Discretization of the Forward and Inverse Problems 116

5.2.1 Forward Problem Discretization 116

5.2.2 Approximation of $\mathcal{T}$ and γ with finite element solutions G_j and G_i^* 117

5.2.3 Discretization of the inverse problem 119

5.2.4 Summary: The inverse problem and its approximations 120

5.3 Discretization-based Error Analysis 121

5.3.1 Effect of forward problem discretization 122

5.3.2 Effect of inverse problem discretization 124

5.3.3 Iterative linearization by Born approximation 126

5.4 Adaptive Mesh Generation 128

5.4.1 Adaptive mesh generation for the forward problem 128

5.4.2 Adaptive mesh generation for the inverse problem 130

5.4.3 Computational complexity of the adaptive mesh generation algorithms 133

5.5 Conclusion 135

6. Conclusion 140

LITERATURE CITED 143

APPENDICES

A. Appendix for chapter 2 152

A.1 Boundedness of $\mathcal{A}_a$ 152

A.2 Compactness of $\mathcal{A}_a$ 153

A.3 Proof of the Lemma 153

B. Appendix for chapter 3 154
  B.1 Solution of the Model Problem (shorter version) 154
  B.2 Solution of the Model Problem 154
    B.2.1 Homogeneous Optical Background 155
    B.2.2 The scattered field due to the circular inclusion 158

C. Appendix for chapter 4 160
  C.1 Proof of theorem 1 160
  C.2 Proof of theorem 2 166
  C.3 Proof of theorem 3 169

D. Appendix 170
  D.1 Finite Element Discretization of the Forward Problem 170
    D.1.1 Existence and Uniqueness of the Solution to the Variational Problem 171
    D.1.2 Finite Element Discretization 174
    D.1.3 Generation of Element Matrices and Vectors 175

LIST OF TABLES

2.1 Definition of variables, functions, and operators. 22

2.2 Definition of function spaces and norms. 23

3.1 Definition of function spaces and norms. 48

3.2 The mesh scenarios and the background μ_a values in simulation study 1. 61

3.3 The error $\|\alpha^\lambda - \tilde{\alpha}_n^\lambda\|_{L^1(\Omega)}$ for each experiment described in the simulation study 1 and Table 3.2. The first column shows the type of the meshes used in the forward and inverse problems, respectively. The unit of background μ_a is cm^{-1}. 63

3.4 The L^1 norm of α^λ and the error $\|\alpha^\lambda - \tilde{\alpha}_n^\lambda\|_{L^1(\Omega)}$ for each experiment described in the simulation study 2. The first column shows the type of the meshes used in the forward and inverse problems, respectively. The radius of the circular heterogeneity is given in *cm*. 66

3.5 The relevant parameters in the experiments 1-5 in simulation study 3. The abbreviation “Conv.” implies that the corresponding mesh was generated using the conventional *a priori* discretization error estimates (3.14)-(3.15) for the forward problem solution, and the conventional *a priori* interpolation error estimate (3.11) for the inverse problem solution. The abbreviation “Prop.” refers to the adaptive meshes generated by using the proposed adaptive mesh generation algorithms based on Theorems 1 and 2, for the inverse and forward problem solutions, respectively. The last column in the table shows the coordinates of the center of the circular heterogeneity, considered in each experiment. 69

3.6 The error $\|\alpha^\lambda - \tilde{\alpha}_n^\lambda\|_{L^1(\Omega)}$ for each experiment described in the simulation study 3. The first column shows the type of the meshes used in the forward and inverse problems, respectively. The superscript ‘C’ denotes that the corresponding adaptive mesh generation is based on the conventional *a priori* error estimates (3.14)-(3.15) and (3.11). 70

4.1 The list of notation. 101

4.2 Definition of function spaces and norms. 102

4.3 The error norm $\|\alpha - \tilde{\alpha}^\lambda\|_0$ measured for each image reconstruction in simulation study 1 (normalized by $\|\alpha - \tilde{\alpha}^\lambda\|_0$ for $\alpha = 0.040$ cm^{-1}.) . . . 102

4.4 The error norm $\|\alpha - \tilde{\alpha}^{\lambda}\|_{L^1(\Omega)}$ measured for each calculation in simulation study 2 (normalized by $\|\alpha - \tilde{\alpha}^{\lambda}\|_0$ for $\alpha = 0.040$ cm^{-1} in the first experiment. 102

5.1 Definition of variables, functions, and operators. 138

5.2 Definition of function spaces and norms. 139

LIST OF FIGURES

1.1 A typical solution of the diffusion equation for a point source. 4

1.2 Sample uniform and adaptive meshes. 6

1.3 An example of crosstalk in simultaneous reconstruction of absorption and diffusion coefficients. 6

1.4 A typical optical setup with transmission geometry with 29 sources (triangles) and 29 detectors. 9

3.1 The setups used for the simulation studies 1, 2, and 3. The squares and triangles denote the detectors and sources, respectively. 60

3.2 Examples of meshes used in the first simulation study. 62

3.3 The reconstruction results of simulation study 1, with the background $\mu_a = 0.032$ cm^{-1}. 64

3.4 The results of simulation study 1, with the background $\mu_a = 0.050$ cm^{-1}. 65

3.5 The cross-sectional views from the reconstructed images in simulation study 1, corresponding to the cases where the background $\mu_a = 0.032$ and $\mu_a = 0.050$ cm^{-1}, respectively. 72

3.6 The results of simulation study 2. The left and right columns show the reconstructed images regarding the optical heterogeneity with radius 0.50 cm, and 1.25 cm, respectively. The background $\mu_a = 0.040$ cm^{-1} in all of the reconstructions. The reference images shown in (a) and (b) are obtained using a uniform mesh with 61 × 61 nodes in both the forward and inverse problems. 73

3.7 **(a)-(b)** Samples of adaptive meshes in the third simulation study, generated by using the conventional error estimates (3.14) and (3.15), which led to unstable optical image reconstruction shown in (c) to (f), for the circular heterogeneity centered at $(3.0, 3.5)$. **(c)-(d)** The unstable optical image reconstructions in the third simulation study, obtained by using the adaptive meshes for the forward problem solution whose examples are shown in (a)-(b). $\lambda = 10^{-8}$. **(e)-(f)** The unstable optical image reconstructions in the third simulation study, obtained by using the adaptive meshes for the forward problem solution whose examples are shown in (a)-(b). λ was set to 10^{-6} to suppress the significantly large artifacts observed in (c)-(d). 74

3.8 Samples of adaptive meshes used in the third simulation study, which led to the optical image reconstructions shown in Figure 3.10. The meshes were generated for the circular heterogeneity centered at (3.0,4.5). 75

3.9 The reconstructed optical images regarding the circular heterogeneity centered at $(3.0, 3.5)$ in the third simulation study. 76

3.10 The reconstructed optical images regarding the circular heterogeneity centered at $(3.0, 4.0)$ in the third simulation study. 77

4.1 The setups used for the simulation studies 1 and 2. The squares and triangles denote the detectors and sources, respectively. 95

4.2 The $L^2(\Omega)$ norm of the scattered field vs each source position (on x-axis) in each of the experiments in the first (a) and second (b) simulation studies. 97

4.3 The cross-sections of $|u_0^j\alpha|$ and $|g_i^*\tilde{\alpha}^\lambda|$ in simulation study 1, which constitute the bound (4.34) for the error resulting from Born approximation. The jth source is located at $(3, 0)$ and the ith adjoint source is located at $(3, 6)$ 103

4.4 Samples of the reconstructed images in the first simulation study. 104

4.5 The cross-sections of $|u_0^j\alpha|$ and $|g_i^*\tilde{\alpha}^\lambda|$ in simulation study 2, which constitute the bound (4.34) for the error resulting from Born approximation. The jth source is located at $(3, 0)$ and the ith adjoint source is located at $(3, 6)$. 105

4.6 Samples of the reconstructed images in the second simulation study. . . 106

5.1 An example of crosstalk in simultaneous reconstruction of absorption and diffusion coefficients. 137

5.2 Samples of the reconstructed images in the second simulation study. . . 139

B.1 The setup considered for the solution of the model problem in equation (B.1). 155

ACKNOWLEDGMENT

Babama (to my Dad)

I would like to dedicate my Ph.D degree to my father, who always knew...

I would like to specially thank my advisor Dr. Birsen Yazıcı for discovering my potential and inviting me for Ph.D program to the USA. I am grateful for her patience, guidance, help, and support during times that I believed getting my Ph.D degree was impossible. I am also thankful to her for providing me with great working conditions and the great research topic and for the opportunity to freely do research under her supervision and guidance.

This is a small step for the mankind but a giant leap for myself. Do not get me wrong; I did pretty good work. I am just trying to be modest. Perhaps, it is not a good time to be modest. All right, this is a small step for me, but a giant leap for the mankind.

But it was difficult. Not everyone should do a Ph.D, nor is there a need for this. I felt that I should have lived the Ph.D education period to do endless research and to feed my need and enthusiasm for learning and knowing more.

I would like to thank my family for their endless support and for repeatedly saying “Don’t worry, you will do fine”. I worried, but I think I did fine. Maybe it is not possible to do fine without worrying.

I would like to thank Dr. Xavier Intes and Dr. Eldar Giladi for their support and help on my research and during my Ph.D life.

I would like to thank Dr. Kiwoon Kwon and Dr. Laurel Reilly for fruitful discussions and for their friendship.

I would like to thank my thesis committee members for their feedback, which helped me make this thesis better.

I would like to thank Dr. Aydan Erkmen, Dr. Gozde Bozdagi, Dr. Murat Eyuboglu, and Dr. Tolga Ciloglu from Middle East Technical University who encouraged me to go to the graduate school.

I would like to thank Dr. Nihat Bilgutay from Drexel University for his support

during the difficult first days and Dr. Dagmar Niebur for her help and encouragement during the wonderful course series at Drexel University.

I would like to thank Dr. Vitor Schneider for his help and friendship during my internship at Corning Inc.

I would like to thank Ann Bruno, Priscilla Magilligan, Laraine Michealides, and George Narode for their patience, interest, and help. I always forgot to do something, and they were there to remind and help me.

I would like to thank the human resources guy who did not hire me to a consulting company back in 2001, suggesting that I should deal with "more technical stuff". I took his advice.

I would like to thank Burak Alaçam. Both of us went through almost the same in the last few years. We shared a lot, we fought a lot; we had fun and suffered together.

I would like to thank my roommates and all of my friends. I have been annoying for a long time and they must have been suffering. I hope everything is going to be better after I get my Ph.D degree.

I would like to thank Pink Floyd for those long hours of music that I listened to during my simulations.

I would like to thank the makers of the show "Modern Sabahlar" (Modern Mornings) on RadyODTU which made it easier to wait for the simulation results (thanks to the 7-hour time difference).

I would like to thank Fenerbahçe for being the greatest soccer team on Earth. They are absolutely the greatest even far far away (Uzaklarda da en büyüksün!). They made the weekends and weekdays more enjoyable.

And thanks to all members of forum antu. It was great to share ideas and opinions about Fenerbahçe, life, and speeding tickets.

Finally, I wish I had written all my daily experience on a regular basis. I would have had a magnificent source for the new graduate students. I even thought of a nice title: "Ph.D or not Ph.D". Maybe I will do it during my next Ph.D degree program.

Ph.D... an adventure.

ABSTRACT

Diffuse Optical Tomography (DOT) is a minimally invasive imaging modality that makes use of the light in the Near-Infrared (NIR) spectrum. The inverse problem in DOT involves reconstruction of spatially varying absorption and diffusion coefficients as well as fluorophore lifetime and yield in tissues from boundary measurements. These fundamental quantities can be utilized to obtain tissue oxy- and deoxyhemoglobin concentrations, blood oxygen saturation, water and fat amount, and to observe uptake and release of contrast agents and organelle concentration in tissue. The unique physiological and biochemical information offered by DOT is very valuable for practical applications such as breast cancer diagnosis, cognitive activity monitoring, brain tumor and hemorrhage detection, functional muscle imaging with a growing list of applications in molecular and cellular imaging.

Diffuse Optical Tomography (DOT) poses a nonlinear ill-posed inverse problem. Furthermore, propagation of NIR light is not restricted to a plane owing to the diffuse nature of photons in turbid media. As a result, DOT is an inherent 3D problem and suffers from low spatial resolution. One has to address all of these drawbacks in order to provide accurate and computationally viable optical image reconstructions.

This thesis focuses on the factors that affect the accuracy of DOT imaging and on how to eliminate these factors. In this context, we present an error analysis to show the effect of the discretization of the forward and inverse problems, and the linearization of the inverse problem on the accuracy of the reconstructed optical images.

First, we consider the inverse problem for which we reconstruct the unknown optical absorption coefficient of a bounded optical medium while the optical diffusion coefficient is assumed to be known. Then, we analyze the error in the reconstructed optical absorption images resulting from the discretization of the forward and inverse problems and the linearization of the inverse problem by Born approximation. Our analysis identifies several factors which influence the extent to which the discretization and Born approximation impact the accuracy of the reconstructed optical absorption images. For example; the mutual dependence of the forward and inverse problems; the number of sources and detectors, their configuration and their locations with respect to optical heterogeneities. Based on the error analysis, we propose novel adaptive discretization schemes for the forward and inverse problems. The proposed discretization schemes lead to adaptively refined composite meshes that yield the desired level of imaging accuracy while reducing the size of the discretized forward and inverse problems.

Finally, we extend our error analysis for the simultaneous reconstruction of the unknown optical absorption and diffusion coefficients. For this problem, while the model for NIR light propagation remains the same, the inverse problem formulation becomes more challenging since two parameters need to be estimated simultaneously. Our analysis shows that the error in the reconstructed optical images due to forward problem discretization depends on the solutions of the inverse problem. Similarly, the error due to inverse problem discretization depends on the solutions of the forward problem, thereby implying the inter-dependence of the forward and inverse problems. One important implication of the analysis is that poor discretization of one optical parameter may lead to error in the reconstruction of the other. Based on the error analysis, we develop adaptive mesh design algorithms which are of low computational complexity as compared to the computational complexity of solving the respective problems, namely the discrete forward and inverse problems.

CHAPTER 1

Introduction

Diffuse Optical Tomography (DOT) is a minimally invasive imaging modality that makes use of the light in the Near-Infrared (NIR) spectrum [87, 43, 42, 49]. The inverse problem in DOT involves reconstruction of spatially varying absorption and scattering properties [72, 15, 4]. These quantities can be used to compute the oxy- and deoxyhemoglobin concentrations, blood oxygen saturation, and the amount of water and fat in tissue [56]. Similarly, diffuse optical fluorescence tomography is concerned with the estimation of fluorophore lifetime and yield in tissues from boundary measurements [23, 41, 29], which can be utilized to observe the uptake and release of contrast agents in tissue [56]. The unique physiological and biochemical information offered by DOT is very valuable for practical applications such as

- breast cancer diagnosis [21, 79, 51],
- cognitive activity monitoring [80, 84, 22],
- brain tumor and hemorrhage detection [42],
- functional muscle imaging [75], and
- molecular and cellular imaging [32, 86].

DOT poses a nonlinear ill-posed inverse problem [4] with a relatively limited number of measurements. Furthermore, the propagation of NIR light is not restricted to a plane owing to the diffuse nature of photons in turbid media. Hence, DOT image reconstruction is an inherent 3D problem, which suffers from poor spatial resolution. In addition, the solution of the inverse problem calls for the solution of the forward (direct) problem. As a result, DOT image reconstruction is an involved mathematical problem with several computational challenges.

In this work, we focus on the computational aspects of the DOT problem. In this context, we examine the factors that affect the accuracy and computational complexity of DOT imaging, such as discretization of the forward and inverse problems and linearization of the inverse problem. Based on the analysis of these factors, we propose algorithms/methods to address the computational complexity of the DOT imaging problem, without compromising the imaging accuracy.

In the following we briefly give an overview of the two aspects of the DOT imaging problem: DOT in a clinical setting and DOT as an imaging problem. The first part of the discussion summarizes the motivation for the use of DOT as a clinical tool. The second part defines briefly the steps in formulating, modeling and solving the DOT imaging problem and describes the challenges therein.

1.1 DOT in a Clinical Perspective

In anatomical imaging, the main interest is the imaging of internal structures. On the other hand, functional imaging is concerned with imaging physiological changes in the body. Examples of anatomical imaging modalities include x-ray imaging, magnetic resonance imaging (MRI) and ultrasound (US). Functional imaging modalities include for example positron emission tomography (PET), single photon emission computed tomography (SPECT), and optical tomography.

In tissue, light is attenuated in the visible region of the spectrum (with wavelength between 450 and 700 nm) and can penetrate less than one centimeter. On the contrary, in the near infrared (NIR) region (wavelengths 700 to 1000 nm), the absorption is significantly lower and light can travel up to 8 centimeters of tissue before it is totally attenuated. Thus, NIR light can be used to interrogate tissue.

DOT offers several potential advantages over existing radiological techniques [42]:

- NIR radiation is minimally invasive, thus can be safely used for screening and continuous monitoring.
- DOT is inexpensive and portable as compared to x-ray computed tomography (CT) and magnetic resonance imaging (MRI).
- DOT offers the potential to differentiate between soft tissues which differ in optical properties such as absorption, scattering coefficients, which are linked to tissue chromophores (water, Haemoglobin (Hb) etc.). Note that more than 97% of the transported oxygen is delivered by oxyhaemoglobin (HbO_2) [58].

Thus, DOT provides information about the metabolic state of an organ by measuring the concentrations of Hb and HbO_2 in the tissue using near-infrared light. On the contrary, MR can not provide this kind of information.

- NIR methods offer novel criteria for cancer differentiation with the ability to in-vivo measure oxygenation and vascularization state, the uptake and release of contrast agents and organelle concentration. These properties are believed to be malignancy specific and may significantly contribute to increased specificity in malignancy diagnosis. Therefore, DOT has the potential as a functional imaging modality, and its applications range from brain oxygenation monitoring in newborn and preterm infants to brain activation studies during physical exercise or mental stimulation, and breast cancer detection/diagnosis and screening.

1.2 Diffuse Optical Tomography as an Imaging Problem

"Solution of an inverse problem entails determining unknown causes based on observation of their effects" (Oleg Mikailivitch Alifanov)... An inverse problem invokes the need for a description of the relationship between these causes and their corresponding effects. This relationship is provided by the forward (direct) problem. Hence, solving an inverse problem calls for a well-defined forward problem and its solution. As a result, DOT imaging requires the definition and solution of a forward and an inverse problem. The forward problem in DOT can be defined by a model that predicts the propagation of the NIR light in a bounded domain given the optical coefficients of the medium. Then, the inverse DOT problem is to estimate the unknown optical coefficients (namely the scattering and absorption coefficients) of the medium from the boundary measurements based on this forward model. In this respect, the forward and inverse problems are the inter-dependent components of the DOT imaging.

In the following, we will describe the components of DOT imaging together with the associated challenges. Note that each of the following can be considered as a factor that affects the accuracy of DOT imaging:

1. **Model for the NIR light propagation (forward problem):** Before image reconstruction can be attempted, a model of photon transport in tissue is required.

   In optical tomography of biological tissues, the use of the radiative transport equation (RTE) is frequently required to accurately describe propagation of multiply scattered light photons. This is especially true in close proximity to sources or boundaries, or in regions with high absorption and low scattering [73]. RTE is a conservation equation which states that the radiance (the number of photons per unit volume), for photons travelling from point r in direction s at time t is equal to the sum of all the mechanisms which increase the radiance minus those effects which reduce it [34]. Provided the scattering coefficient is much larger than the absorption coefficient, the diffusion approximation to RTE is a commonly used model, especially in DOT imaging of deep tissue such as breast (see in figure (1) for a typical solution of the frequency-domain diffusion equation on a bounded domain). We refer to [4] for a more detailed review.

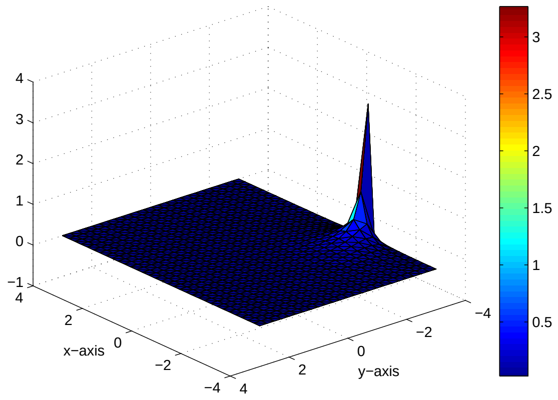


Figure 1.1: A typical solution of the diffusion equation for a point source.

2. **Ill-posedness of the inverse problem**: While the forward problem is well-defined and well-posed, the inverse problem in DOT is ill-posed. The ill-posedness (in Hadamard sense [52]) in general implies:

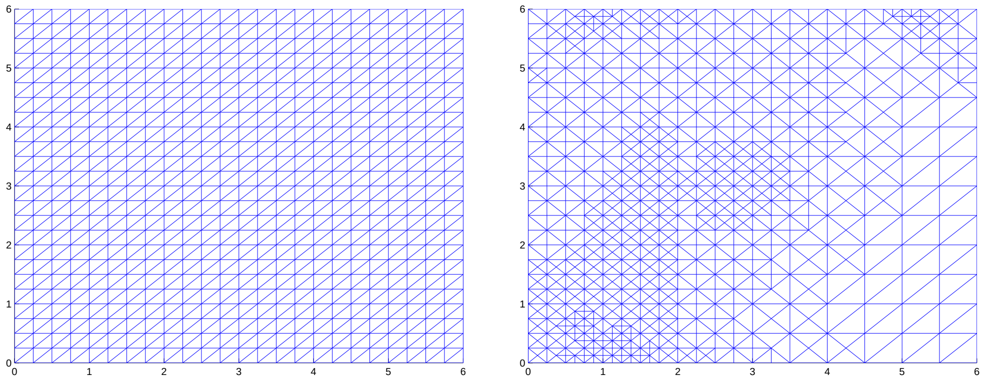
- The solution does not depend continuously on the measurements (boundary data).
- There is no unique solution to the inverse problem.
- There is no solution to the inverse problem.

In DOT, a solution exists provided there is sufficient amount of data. On the other hand, the solution is not unique [6] owing to the null-space of the operator which relates the unknown coefficients to the boundary data. Furthermore, the solution is not stable, i.e. a small variation in the measurement data can lead to large deviations in the solution.

3. **Discretization and size of the discretized forward and inverse problems:** The propagation of the photons in biological tissues is isotropic due to large and isotropic scattering coefficient of the medium. This makes the DOT inverse problem an inherent 3D problem.

   In general, no analytical solutions are available for the forward problem. Thus, analytical solutions are rare to obtain for the inverse problem as well. In this context, one can obtain only a finite-dimensional approximation of the solutions to the forward and inverse problems. The size of the forward and inverse problems may become large, extent of which depends on the discretization of these problems. We note that poor discretization of the forward and inverse problems may lead to severe artifacts in the reconstructed optical images (see chapters 2-3-4). See figures (1.2(a)) and (1.2) for sample two dimensional uniform and adaptive meshes with triangular elements.

4. **Dual-parameter:** The inverse problem in optical tomography is concerned with reconstructing the optical coefficients of the medium under inspection. In general, one has to reconstruct both the absorption and scattering coefficients of a medium simultaneously. As a result, the inverse problem becomes larger in size with the same number of measurements, which implies a more challenging inverse problem with an increase in its computational complexity. Furthermore, dual-parameter reconstruction often ends up with the so-called

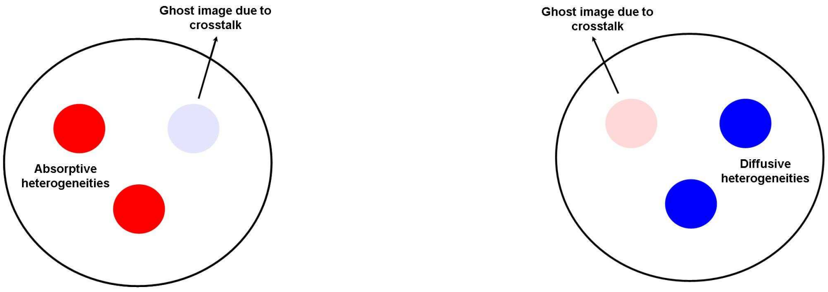


(a) A sample uniform coarse mesh that can be used for the discretization of the forward or inverse problem.

(b) An adaptive mesh that can be used for the discretization of the forward or inverse problem.

Figure 1.2: Sample uniform and adaptive meshes.

crosstalk problem between the two images [65, 74, 26, 44], which degrades the imaging accuracy (see figures (1.3(a)) and (1.3(b)) for an illustration of crosstalk).



(a) Reconstructed absorption image with crosstalk. Red circles indicate the reconstructed absorptive heterogeneities and the light blue circle on the top right is the ghost image introduced as a result of crosstalk from the diffusion image.

(b) Reconstructed diffusion image with crosstalk. Blue circles indicate the reconstructed absorptive heterogeneities and the pink circle on the top left is the ghost image introduced as a result of crosstalk from the absorption image.

Figure 1.3: An example of crosstalk in simultaneous reconstruction of absorption and diffusion coefficients.

5. **Non-linearity of the inverse problem:** Except for special geometries, there is, in general, no closed form relationship between the unknown optical coefficients and the boundary data. Furthermore, this relationship is nonlinear

including the special cases, in which closed-form solutions are available. This precludes the use of analytic methods such as back-projection methods. Linearization of this relationship (that is linearization of the inverse problem) affords a great simplification at the expense of a compromise in imaging accuracy (see Chapter 4). In this context, Born and Rytov approximation schemes are the widely used approximation schemes [17, 24, 68].

6. **Model mismatch:** Apart from the selection of the NIR light propagation model, model mismatches occur in the presence of measurement noise. Due to the ill-posed nature of the inverse problem, measurement noise can be the largest error source in a DOT imaging system.

   Note that linearization of the inverse problem can also be viewed as re-modeling, which poses a model mismatch problem.

   It is also common to reconstruct only the absorption coefficient of a medium, as it is the main optical property that relates directly to the chromophore concentrations in the medium, which can be used to assess the tumor characteristics. However, assuming a known spatially homogeneous diffusion coefficient for the medium of interest may lead to model mismatches as well. A similar situation is observed in fluorescence tomography when the inverse problem is linearized around a background with homogeneous absorption coefficient at the excitation wavelength [69].

7. **Choice of algorithms and numerical solvers:** Since, in general, no analytical solution is available for the forward and inverse problem, numerical algorithms have to be used. Due to the large size of these problems, and the ill-posedness of the inverse problem, the choice of the numerical solvers and the design of the algorithms becomes important. The convergence rate and the computational complexity are the key factors in the selection of a numerical solver. The computational complexity of a numerical algorithm increases with the size of the discrete forward and inverse problems. Thus, the efficiency of the numerical solver depends on the discretization of the forward and inverse problems, as well.

8. **Inverse problem formulation:** The formulation of the inverse DOT problem has to address the ill-posedness of the inverse problem. This is usually done by using regularizers, which in general accounts for both the non-uniqueness and the instability of the solution in the presence of measurement noise. On the other hand, using regularizers often impose a bias on the reconstructed images. Thus, the choice of regularizers has to be addressed quantitatively (see [37] and the references therein).

9. **Low spatial resolution:** The reason for the low spatial resolution of DOT images is the high scattering coefficient of biological tissues for the visible and near-infrared light, which makes light propagation a diffuse process. Each detected photon undergoes multiple scattering events and deviates from the direct line of sight between the source and detector sites, thereby blurring the spatial information content it carries.

   Both the inherent low resolution and the ill-posedness of the problem make it desirable to combine DOT with a second independent imaging modality such as MRI whose high-resolution spatial information could be used to guide the reconstruction process. In this context, *a priori* anatomical and/or physiological information can be used to assist the design of regularizers which result in minimum bias and variance. Use of *a priori* information may enhance the imaging accuracy and the spatial resolution significantly (see [37] and the references therein).

10. **Information content provided by the imaging setup:** The number of sources and detectors and their orientation with respect to each other and with respect to internal optical heterogeneities determine the size (dimension) of the space of observable objects (the space of functions which can be resolved from the boundary data) in a diffuse optical tomography system [18, 2]. We refer to figure (10) for a sample optical setup with transmission geometry.

    A priori information about the optical properties of a medium can be used to maximize the information content by placing the sources and detectors accordingly or even to compensate for the missing information content. Furthermore, *a priori* information can assist the discretization of the forward and inverse problems (see for example Theorems 1-2 in Chapter 2).

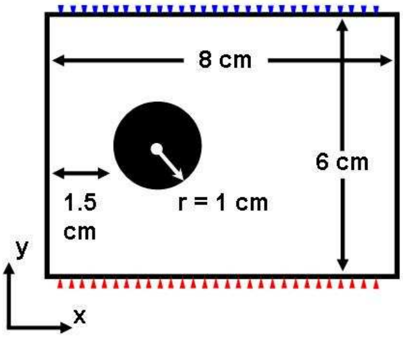


Figure 1.4: A typical optical setup with transmission geometry with 29 sources (triangles) and 29 detectors.

1.3 Thesis Outline

In Section 1.1, we have outlined the motivations for DOT imaging in a clinical setting and in Section 1.2, we have summarized the main components of DOT imaging. In this thesis, we focus on the computational aspects of the DOT imaging problem, which can be identified by the inter-dependent factors described in section 1.2. In particular, we focus on the discretization of the forward and inverse problems and linearization of the inverse problem by Born approximation. We show the factors that determine the extent to which the discretization of the forward and inverse problems and linearization of the inverse problem affect the accuracy of DOT imaging. Note that as we concentrate on the discretization of the forward and inverse problems and linearization of the inverse problem, we discuss and address each of the computational challenges (Section 1.2) associated with DOT imaging.

Below, we present the details of our approach in analyzing the effect of discretization of the forward and inverse problems, and linearization of the inverse problem on the accuracy of DOT imaging.

1.3.1 Effect of discretization on the accuracy of optical absorption imaging

Numerical approaches in solving the forward and inverse problems in DOT poses a tradeoff between computational efficiency and imaging accuracy. This tradeoff is a direct consequence of the discretization of the forward and inverse problems and the size of the resulting discrete forward and inverse problems. The imaging accuracy depends on the discretization error in the forward and inverse problem solutions. On the other hand, attempting to minimize the discretization error in the solutions of both problems separately implies a significant increase in the size of the discrete forward and inverse problems. Hence, it is important to understand the relationship between the discretization error and the resulting error in the solution of the inverse problem. Such a relationship can illuminate the mutual dependence of the forward and inverse problem solutions and identify the factors that control the extent to which the discretization error in the solutions of the forward and inverse problems affects the accuracy.

In Chapter 2 of this work, we present an analysis to show the effect of discretization on the accuracy of DOT imaging. The objectives of this part of the thesis are:

1. To show explicitly the effect of discretization of the forward and inverse problems on the accuracy of the reconstructed optical absorption images.

2. To identify the parameters that influence the extent to which the choice of basis functions and discretization points for the *forward problem discretization* affects the accuracy of the reconstructed optical absorption images.

3. To identify the parameters that influence the extent to which the choice of basis functions and discretization points for the *inverse problem discretization* affects the accuracy of the reconstructed optical absorption images.

4. To show that an effective discretization approach should take into account the inter-dependence of the forward and inverse problem solutions.

5. To obtain a measure of the error resulting from discretization which can be

used in the design and development of adaptive discretization methods for the forward and inverse problems.

To meet these objectives, we first define the forward and inverse problems. In this work, we model the forward problem by the frequency-domain diffusion equation. For the inverse problem, we focus on the estimation of the absorption coefficient. We consider the linear integral equation resulting from the iterative linearization of the inverse problem based on Born approximation and use zeroth order Tikhonov regularization to address the ill-posedness of the resulting integral equation. We use finite elements with first order Lagrange basis functions to discretize the forward and inverse problems and analyze the effect of the discretization of each problem on the reconstructed optical absorption image. In our analysis, we first consider the impact of the inverse problem discretization when there is no discretization error in the solution of the forward problem, and provide a bound for the resulting error in the reconstructed optical image. Next, we analyze the effect of the forward problem discretization on the accuracy of the reconstructed image without discretizing the inverse problem, and obtain another bound for the resulting error in the reconstructed optical image.

As a result, in chapter 2, we summarize the implications of the error analysis in two theorems which provide an insight into the impact of forward and inverse problem discretizations on the accuracy of the reconstructed optical absorption images. These theorems show that the error in the reconstructed optical image due to the discretization of each problem is bounded by roughly the multiplication of the discretization error in the corresponding solution and the solution of the other problem. In particular, the study shows that solving the forward problem (diffusion equation and the associated adjoint problem) accurately may not avoid errors in the reconstructed optical images. Similarly, relatively large discretization error in the solution of the forward problem may have relatively low impact on the accuracy of the reconstructed optical images, depending on the source-detector configuration, and orientation with respect to the optical heterogeneities. Finally, we extend our analysis to show the effect of noise on the accuracy of the reconstructed optical images. Our analysis shows that the presence of noise results in error terms in addition to the error in the reconstructed optical images induced by the discretization of the forward and inverse problems.

The presented analysis brings several advantages:

- A clear relationship between the error in the reconstructed optical absorption images resulting from discretization and the parameters associated with the forward and inverse problems, such as source-detector configuration, orientation of the sources and detectors with respect to the support of the optical absorption coefficient.
- Two new error estimates which are used to develop new adaptive discretization algorithms (See Chapter 3).
- Provided there is *a priori* information about the support of the absorptive heterogeneity, the error estimates can be used to place the sources and detectors such that the error resulting from discretization is reduced.
- The error analysis can be extended to show the effect of the discretization error on the accuracy of the simultaneous reconstruction of scattering and absorption coefficients (See chapter 5).
- The framework used to analyze the error in the reconstructed optical images resulting from discretization can be employed to analyze the error resulting from linearization of the inverse problem by Born approximation (See chapter 4).
- The error analysis introduced in this work is not limited to DOT, and can easily be adapted for similar inverse parameter estimation problems such as electrical impedance tomography, bioluminescence tomography, optical fluorescence tomography, microwave imaging etc.

1.3.2 Adaptive mesh generation for DOT

Based on the two error bounds provided by the error analysis in Chapter 2, we introduce an adaptive discretization scheme for the forward and inverse problems, respectively. The presented adaptive mesh generation algorithms address the interdependence between the solutions of the forward and inverse problems and take into account the orientation of the source-detectors and the absorptive perturbations. This makes the adaptive discretization algorithms introduced in this work different from the previous approaches [28, 36, 82, 35, 46, 54], which make use of conventional error estimates.

The objectives of this study can be summarized as follows:

1. To develop adaptive mesh generation algorithms for the discretization of the forward and inverse problems in DOT, by using the error estimates derived in Chapter 2.

2. To address the tradeoff between the computational efficiency and accuracy in DOT imaging, i.e. to reduce the number of unknowns in the discrete forward and inverse problems, yet sustain the accuracy.

3. To improve spatial resolution of the reconstructed optical absorption images without compromising the computational efficiency.

Our simulation experiments validate the implications of the error analysis presented in Chapter 2 and show that the proposed mesh generation algorithms significantly improve the accuracy of the reconstructed optical images while keeping the number of unknowns in the discrete forward and inverse problems for a given number of unknowns in the discrete forward and inverse problems. We specifically show that using the discretization error estimates, which do not take into account the interdependence of forward and inverse problems as a criterion for discretization, may lead to severely degraded image reconstructions (see simulation study 3 in Chapter 3). We also discuss the computational complexity of the proposed adaptive mesh generation algorithms and compare it to the computational complexity of mesh generation algorithms based on the conventional discretization error estimates.

The advantages of this study are as follows:

- The proposed mesh generation algorithms significantly improve the accuracy of the reconstructed optical images for a given number of unknowns in the discrete forward and inverse problems while conventional error estimates do not include domain specific factors.

- The proposed adaptive mesh generation algorithms can be adapted for similar inverse parameter estimation problems, such as electrical impedance tomography, optical fluorescence tomography, bioluminescence tomography, microwave imaging, etc.

1.3.3 Effect of linearization by Born approximation on the accuracy of DOT imaging

The nonlinearity of the associated inverse problem makes DOT imaging a computationally intense problem, which calls for the use of nonlinear inversion methods. In general, nonlinear inversion algorithms require repetitive solution of the forward problem defined by the light propagation model, which is followed by the update of the inverse problem solution. As a result, nonlinear algorithms provide accuracy at the expense of high computational complexity. A widely used approach to overcome the difficulties posed by the nonlinearity is to linearize the inverse problem using Born (or Rytov) approximation and solve the resulting linear problem to reconstruct the optical parameters [7]. Such approximation schemes assume a small perturbation on the background optical coefficients, which are in general assumed to be spatially homogeneous. Spatially homogeneous backgrounds allow for analytical solutions of the forward problem for certain domain geometries [64] and analytical solutions for the linearized inverse problem as well [63]. Hence, linearization of the inverse DOT problem brings several computational advantages. As a result, Born approximation can be considered as a compromise between accuracy and computational complexity. Intuitively, it can be deduced that the accuracy of the Born approximation is determined by the perturbation in the optical coefficients with respect to a background with known optical properties. In this work, we show what other factors lead to an increase in the error in the reconstructed optical absorption images resulting from linearization by Born approximation.

The objectives of this study can be summarized as follows:

1. To provide an insight into the effect of linearization by Born approximation

on the accuracy of DOT imaging.

2. To present an error analysis to show the parameters that determine the extent to which the linearization by Born approximation affects the accuracy of DOT imaging.
3. To show under what conditions the linearization of the inverse problem by Born approximation may ensure the accuracy of DOT imaging.

Our analysis provides a tight error bound that takes into account the spatial orientation of the optical absorptive heterogeneity, the optical field generated by the light sources and the detector positions. This makes the derived error bound significant, especially in optical media with relatively small sizes.

The advantages of this study and its results are as follows:

- The presented error analysis shows the factors that determine the extent to which the accuracy of DOT imaging is affected by the linearization of the inverse problem.
- The error analysis provides a measure for the choice of step length in iterative linearization based nonlinear optimization methods, such as trust-region algorithms.
- Provided there is *a priori* information about the support of the heterogeneity, the analysis can be used to place the sources and detectors such that the error in the reconstructed images resulting from linearization is reduced. One posterior approach can be to use only the specific sources-detectors in the image reconstruction, such that the error due to linearization by Born approximation is reduced.

1.3.4 Effect of discretization on the simultaneous reconstruction of optical absorption and diffusion coefficients

In Chapter 5 of this work, we present an analysis to show the effect of discretization on the accuracy of simultaneously reconstructed optical absorption and diffusion images. In this respect, we extend our work presented in chapter 2 and 3 for the simultaneous reconstruction of optical absorption and diffusion coefficients. While the light propagation model remains the same as in chapters 2 and 3, the inverse problem formulation needs to be extended to estimate the unknown optical diffusion coefficient as well. Since this is a two-parameter inverse problem, the discretization of the inverse problem implies an increase in the number of unknowns in the resulting discrete inverse problem. Consequently, the tradeoff between computational complexity and accuracy of DOT imaging becomes more critical: One has to minimize the number of unknowns in the discrete inverse problem while providing the desired level of imaging accuracy.

The objectives of this part of the thesis are:

1. To show explicitly the effect of discretization of the forward and inverse problems on the accuracy of simultaneously reconstructed optical absorption and diffusion coefficient images.
2. To identify the parameters that enhance the effect of poor discretization on the accuracy of simultaneous DOT imaging.
3. To show that an effective discretization approach should take into account the inter-dependence of forward and inverse problem solutions.
4. To reveal that the accuracy of one optical parameter reconstruction depends not only on how well it is discretized but also on how well the other optical parameter is discretized, thereby providing a different perspective on and addressing so-called the “inter-parameter cross-talk” resulting from discretization.
5. To obtain an estimate of the error resulting from discretization which can be used in the design and development of adaptive discretization methods for the forward and inverse problems, with relatively low computational complexity as compared to the computational complexity of solving the resulting discrete problems.

In this study, we use an approach similar to the one followed in chapters 2 and 4. In this respect, we consider the iterative linearization of the inverse problem based on Born approximation. In order to regularize the resulting linear integral equation, we use zeroth order Tikhonov regularization. We use finite elements with first order Lagrange basis functions to discretize the forward and inverse problems and analyze the effect of the discretization of each problem on the reconstructed optical absorption and diffusion images. Different than chapter 2, in the analysis we first consider the impact of the forward problem discretization and provide a bound for the resulting error in the reconstructed optical images. Next, we analyze the effect of the discretization of the inverse problem whose formulation is based on the numerical (finite element) solutions of the forward problem, which possess error due to forward problem discretization.

As a result, in chapter 5, we summarize the implications of the error analysis in two theorems which provide an insight into the impact of forward and inverse problem discretizations on the accuracy of the simultaneously reconstructed optical absorption and diffusion images. These theorems show that the error in the reconstructed optical image due to the discretization of each problem is bounded by roughly the multiplication of the discretization error in the corresponding solution and the solution of the other problem. In addition, the analysis of the error due inverse problem discretization reveals that the discretization of the optical diffusion coefficient has an influence on the accuracy of the reconstructed optical absorption coefficient. As a result, one has to take into account the dependence of the error in the reconstructed optical images due to discretization on all of the fundamental elements of DOT imaging; i.e. the solutions of the forward problem for each source and detector, the absorption coefficient, and the diffusion coefficient. Finally, we use the error bounds obtained in the analysis to devise adaptive mesh generation algorithms that can address the compromise between computational efficiency and accuracy. We present two practical adaptive mesh generation algorithms whose computational complexity is relatively low as compared to the computational complexity of solving the resulting discrete forward and inverse problems.

The error analysis and the proposed adaptive mesh generation algorithms provide several advantages including the following:

- The error analysis helps to understand the factors that control the extent to which the discretization affects the accuracy of DOT imaging when both the scattering and absorption coefficients are reconstructed.

- The analysis reveals how the discretization of one optical coefficient will affect the accuracy of the reconstruction of the other optical coefficient. In this respect, the error analysis provides a perspective to understand the inter-parameter cross-talk (see Chapter 5) between the optical absorption and scattering coefficients resulting from discretization.

- The error analysis indicates that an effective discretization scheme for both the forward and inverse problems has to take into account the inter-dependence of the forward and inverse problems. In other words, the discretization of one problem can not be independent of the solution of the other problem.

- The error estimates obtained in the analysis used to develop new discretization algorithms for the simultaneous reconstruction of the optical coefficients. The computational cost of the algorithms does not exceed that of solving the resulting discrete forward and inverse problems.

- The adaptive mesh generation algorithms can be applied *a priori* or *a posteriori* depending on the availability of *a priori* information about the solutions of the forward and inverse problems that are required to compute the error bounds.

- The error analysis and adaptive mesh generation algorithms introduced in this work is not limited to DOT, and can easily be adapted for similar two-parameter inverse problems. In addition, the analysis framework can be used to examine the inter-parameter crosstalk occurring due to other factors, such as linearization by Born approximation, modeling errors, regularization etc.

CHAPTER 2

Effect of discretization error in diffuse optical absorption imaging

2.1 Introduction

Imaging in Diffuse Optical Tomography (DOT) is comprised of two interdependent stages which seek solutions to the forward and inverse problems. The forward problem is associated with describing the Near Infrared (NIR) light propagation, while the objective of the inverse problem is to estimate the unknown optical parameters from boundary measurements [4].

There are a variety of factors that affect the accuracy of the DOT imaging, such as model mismatch (due to light propagation model and/or linearization of the inverse problem), measurement noise, discretization, numerical algorithm efficiency, and inverse problem formulation. In this two-part study, we focus on the effect of discretization of the forward and inverse problems. In the first part of our work, we present an error analysis to show the effect of discretization on the accuracy of the reconstructed optical absorption images. We identify the factors specific to the imaging problem, which determine the extent to which the discretization impacts the accuracy of the reconstructed optical absorption images. In the sequel, first, we use the error analysis to develop novel adaptive discretization algorithms for the forward and inverse problems to reduce the error in the reconstructed optical images resulting from discretization. Next, we present numerical experiments that support the main results of part I and demonstrate the effectiveness of the developed adaptive mesh generation algorithms.

There has been extensive research on the estimation of discretization error in the solutions of partial differential equations (PDEs) [3, 10, 11, 13, 81, 83]. On the contrary, relatively little has been published in the area of parameter estimation problems governed by PDEs. See for example [14] for an *a posteriori* error estimate for the Lagrangian in the inverse scattering problem for the time-dependent acoustic wave equation and [62] for *a posteriori* error estimates for distributed elliptic optimal control problems. In the area of DOT, it was numerically shown that the approximation errors resulting from the discretization of the forward problem can lead to significant errors in the reconstructed optical images [5]. However, an analysis regarding the error in the reconstructed optical images resulting from discretization has not been reported so far.

In this work, we model the forward problem by the frequency-domain diffusion equation. For the inverse problem, we focus on the estimation of the absorption coefficient. We consider the linear integral equation resulting from the iterative linearization of the inverse problem based on Born approximation and use zeroth order Tikhonov regularization to address the ill-posedness of the resulting integral equation. We use finite elements with first order Lagrange basis functions to discretize the forward and inverse problems and analyze the effect of the discretization of each problem on the reconstructed optical absorption image. Our analysis describes the dependence of the image quality on the optical image properties, the configuration of the source and detectors, the orientation of the source and detectors with respect to absorptive heterogeneities, and on the regularization parameter in addition to the discretization error in the solution of each problem. In our analysis, we first consider the impact of the inverse problem discretization when there is no discretization error in the solution of the forward problem, and provide a bound for the resulting error in the reconstructed optical image. Next, we analyze the effect of the forward problem discretization on the accuracy of the reconstructed image without discretizing the inverse problem, and obtain another bound for the resulting error in the reconstructed optical image. We see that each error bound comprises the discretization error in the corresponding problem solution, scaled spatially by the solutions of both problems. This is a direct consequence of the fact that the inverse problem solution depends on the model defined by the forward problem. As a result, the error analysis yields specific error estimates which are different than the conventional discretization error estimates (see equations (5.25)-(5.26) and (2.41)) which take into account only the smoothness and support of the function of interest, and the finite dimensional space of approximating functions [20]. We further discuss the use of other basis functions and methods in the discretization of the forward and inverse problems and explain how the error bounds can be modified accordingly. Finally, we extend our analysis to show the effect of noise on the accuracy of the reconstructed optical images. Our analysis shows that the presence of noise results in error terms in addition to the error in the reconstructed optical images induced by the discretization of the forward and inverse problems.

This work not only provides an insight into the error in reconstructed optical absorption images resulting from discretization, but also motivates the development of novel adaptive mesh generation algorithms to address this error [40]. In addition, the analysis presented in this work provides a means to identify and analyze the error in the reconstructed optical images resulting from the linearization of the Lippmann-Schwinger type equations [25] using Born approximation [38]. Furthermore, the error analysis introduced in this paper is not limited to DOT, and can easily be extended for use in similar inverse parameter estimation problems such as electrical impedance tomography, bioluminescence tomography, optical fluorescence tomography, microwave imaging etc., in all of which the inverse problem can be interpreted in terms of a linear integral equation, whose kernel is the solution of a PDE that models the forward problem.

The outline of this paper is as follows: Section 2 defines the forward and inverse problems. In Section 3, we discuss the discretization of the forward and inverse problems. In Section 4, we present two theorems that summarize the impact of discretization on the accuracy of the reconstructed optical images, which is followed by the Conclusion section. The Appendix includes results regarding the boundedness and compactness of the linear integral operator used to define the inverse problem, and the proof for the convergence of the inverse problem discretization.

2.2 Forward and Inverse Problems

In this section, we describe the model for NIR light propagation and define the forward and inverse DOT problems. Table 5.1 provides a list of the notation and Table 5.2 provides the definition of function spaces and norms used throughout the paper. We note that we use calligraphic letters to denote the operators, e.g. $\mathcal{A}_a$, $\mathcal{I}$, $\mathcal{K}$ etc.

Table 2.1: **Definition of variables, functions, and operators.**

Notation	Explanation
Ω	Bounded domain in $\mathbb{R}^3$ with Lipschitz boundary
$\partial\Omega$	Lipschitz boundary of Ω
$\mathbf{x}$	Position vector in $\Omega \cup \partial\Omega$
$g_j(\mathbf{x})$	Solution of the diffusion equation at $\mathbf{x}$ for the j^{th} point source located at $\mathbf{x}_s^j$
$g_i^*(\mathbf{x})$	Solution of the adjoint problem at $\mathbf{x}$ for the i^{th} adjoint source located at $\mathbf{x}_d^i$
$G_j(\mathbf{x})$	Finite element approximation of g_j at $\mathbf{x}$
$G_i^*(\mathbf{x})$	Finite element approximation of g_i^* at $\mathbf{x}$
$e_j(\mathbf{x})$	The discretization error at $\mathbf{x}$ in the finite element approximation of g_j
$e_i^*(\mathbf{x})$	The discretization error at $\mathbf{x}$ in the finite element approximation of g_i^*
$\alpha(\mathbf{x})$	Small perturbation over the background μ_a at $\mathbf{x}$
$\Gamma_{i,j}$	Differential measurement at the i^{th} detector due to the j^{th} source
$\mathcal{A}_a$	The integral operator mapping $\alpha \in L^\infty(\Omega)$ to $\Gamma_{i,j} \in \mathbb{C}^{N_d \times N_s}$
$\mathcal{A}_a^*$	The adjoint of $\mathcal{A}_a$ mapping from $\mathbb{C}^{N_d \times N_s}$ to $L^1(\Omega)$
$H_{i,j}(\mathbf{x})$	The kernel of $\mathcal{A}_a$ at $\mathbf{x}$
$H_{i,j}^*(\mathbf{x})$	The kernel of $\mathcal{A}_a^*$ at $\mathbf{x}$
$\gamma(\mathbf{x})$	$\mathcal{A}_a^*\Gamma_{i,j}$ at $\mathbf{x}$
λ	The regularization parameter
$\alpha^\lambda(\mathbf{x})$	Solution of the regularized inverse problem at $\mathbf{x}$
$\alpha_n^\lambda(\mathbf{x})$	Solution of the discretized regularized inverse problem with exact kernel at $\mathbf{x}$
$\tilde{\alpha}^\lambda(\mathbf{x})$	Solution of the regularized inverse problem with degenerate kernel at $\mathbf{x}$
$\tilde{\alpha}_n^\lambda(\mathbf{x})$	Solution of the discretized regularized inverse problem with degenerate kernel at $\mathbf{x}$

2.2.1 Forward Problem

We use the following boundary value problem to model the NIR light propagation in a bounded domain $\Omega \subset \mathbb{R}^3$ with Lipschitz boundary $\partial\Omega$ [20, 4]:

$$-\nabla \cdot D(\mathbf{x})\nabla g_j(\mathbf{x}) + \left(\mu_a(\mathbf{x}) + \frac{i\omega}{c}\right) g_j(\mathbf{x}) = Q_j(\mathbf{x}) \quad \mathbf{x} \in \Omega, \tag{2.1}$$

$$g_j(\mathbf{x}) + 2aD(\mathbf{x})\frac{\partial g_j}{\partial n}(\mathbf{x}) = 0 \quad \mathbf{x} \in \partial\Omega, \tag{2.2}$$

Table 2.2: **Definition of function spaces and norms.**

Notation	Explanation
$\overline{f}$	The complex conjugate of the function f
$C(\Omega)$	Space of continuous complex-valued functions on Ω
$C^k(\Omega)$	Space of complex-valued k-times continuously differentiable functions on Ω
$L^\infty(\Omega)$	$L^\infty(\Omega) = \{f \mid ess\,\sup_\Omega \lvert f(\mathbf{x})\rvert < \infty\ \}$
$L^p(\Omega)$	$L^p(\Omega) = \{f \mid (\int_\Omega \lvert f(\mathbf{x})\rvert^p d\mathbf{x})^{1/p} < \infty\ \}$, $p \in [1,\infty)$
$D_w^z f$	z^{th} weak derivative of f
$H^p(\Omega)$	$H^p(\Omega) = \{f \mid (\sum_{\lvert z\rvert \le p} \lVert D_w^z f\rVert_0^2)^{1/2} < \infty\ \}$, $p \in [1,\infty)$
$\lVert f\rVert_0$	The $L^2(\Omega)$ norm of f
$\lVert f\rVert_p$	The $H^p(\Omega)$ norm of f
$\lVert f\rVert_\infty$	The $L^\infty(\Omega)$ norm of f
$\lVert f\rVert_{L^p(\Omega)}$	The $L^p(\Omega)$ norm of f
$\lVert f\rVert_{0,m}$	The L^2 norm of f over the m^{th} finite element Ω_m
$\lVert f\rVert_{p,m}$	The H^p norm of f over the m^{th} finite element Ω_m

where $g_j(\mathbf{x})$ is the photon density at $\mathbf{x}$, Q_j is the point source located at $\mathbf{x}_s^j$, $D(\mathbf{x})$ is the diffusion coefficient and $\mu_a(\mathbf{x})$ is the absorption coefficient at $\mathbf{x}$, $i = \sqrt{-1}$, ω is the modulation frequency of the source, c is the speed of the light, $a = (1+R)/(1-R)$ where R is a parameter governing the internal reflection at the boundary $\partial\Omega$, and $\partial \cdot /\partial n$ denotes the directional derivative along the unit normal vector on the boundary. Note that we assume the diffusion coefficient is isotropic. For the general anisotropic material, see [55].

The adjoint problem [4] associated with (5.1)-(5.2) is given by the following boundary value problem:

$$-\nabla \cdot D(\mathbf{x})\nabla g_i^*(\mathbf{x}) + \left(\mu_a(\mathbf{x}) - \frac{i\omega}{c}\right) g_i^*(\mathbf{x}) = 0 \quad \mathbf{x} \in \Omega, \tag{2.3}$$

$$g_i^*(\mathbf{x}) + 2aD(\mathbf{x})\frac{\partial g_i^*}{\partial n}(\mathbf{x}) = Q_i^*(\mathbf{x}) \quad \mathbf{x} \in \partial\Omega, \tag{2.4}$$

where Q_i^* is the adjoint source located at $\mathbf{x}_d^i$. We note that we approximate the point source Q_j in (5.1) and the adjoint source Q_i^* in (5.4) by Gaussian functions with sufficiently low variance, whose centers are located at $\mathbf{x}_s^j$ and $\mathbf{x}_d^i$, respectively.

In this work, we consider the finite-element approximations of the solutions of the forward problem. Hence, before we discretize the forward problem (see Section 5.2.1), we consider the variational formulations of (5.1)-(5.2) and (5.3)-(5.4) by multiplying (5.1) by a test function $\overline{\phi} \in H^1(\Omega)$ and integrating over Ω [20]:

$$\int_{\Omega} \left[\nabla\overline{\phi} \cdot D\nabla g_j + \overline{\phi}\left(\mu_a + \frac{i\omega}{c}\right) g_j - \overline{\phi} Q_j \right] d\mathbf{x} + \frac{1}{2a}\int_{\partial\Omega} \overline{\phi} g_j dl \;\; = \;\; 0, \qquad (2.5)$$

where the boundary integral term results from the boundary condition (5.2).

Equivalently, we can express (D.6) by defining the sesquilinear form $b(\phi, g_j)$[47]:

$$b(\phi, g_j) := A(\phi, g_j) + \left\langle \phi, \frac{1}{2a} g_j \right\rangle = (\phi, Q_j), \qquad (2.6)$$

where

$$\begin{aligned} A(\phi, g_j) \;&:=\; \int_{\Omega} \left[\nabla\overline{\phi} \cdot D\nabla g_j + \left(\mu_a + \frac{i\omega}{c}\right) \overline{\phi} g_j \right] d\mathbf{x}, \\ (\phi, Q_j) \;&:=\; \int_{\Omega} \overline{\phi} Q_j d\mathbf{x}, \\ \left\langle \phi, \frac{1}{2a} g_j \right\rangle \;&:=\; \frac{1}{2a}\int_{\partial\Omega} \overline{\phi} g_j dl. \end{aligned}$$

Similarly, the variational problem for (5.3)-(5.4) can be formulated by defining the sesquilinear form $b^*(\phi, g_i^*)$:

$$b^*(\phi, g_i^*) := A(\phi, g_i^*) + \left\langle \phi, \frac{1}{2a} g_i^* \right\rangle = \left\langle \phi, \frac{1}{2a} Q_i^* \right\rangle, \qquad (2.7)$$

where in $A(\phi, g_i^*)$, ω is replaced by $-\omega$.

The sesquilinear forms $b(\phi, g_j)$, $b^*(\phi, g_i^*)$ are continuous and positive definite for bounded D and μ_a [47]. As a result, the variational problems (5.6) and (5.7) have unique solutions, which follows from the Lax-Milgram Lemma [20]. The solutions g_j and g_i^* of the variational problems (5.6) and (5.7) belong to $H^1(\Omega)$, which results from the H^1-boundedness of the Gaussian function that approximates the point source Q_j and the adjoint source Q_i^* [47]. Assuming $D, \mu_a \in C^1(\Omega)$ and noting that $Q_j, Q_i^* \in H^1(\Omega)$; the solutions g_j, g_i^* satisfy $g_j, g_i^* \in H^2_{loc}(\Omega)$ (Chapter 6.3, Theorem

2 in [30]). This last condition implies (Chapter 5.6, Theorem 6 in [30])

$$g_j, g_i^* \in C(\Omega). \tag{2.8}$$

2.2.2 Inverse Problem

In this work, we focus on the estimation of the absorption coefficient; therefore we assume $D(\mathbf{x})$ is known for all $\mathbf{x} \in \Omega \cup \partial\Omega$. To address the nonlinear nature of the inverse DOT problem, we consider an iterative algorithm based on repetitive linearization of the inverse problem using first order Born approximation [4]. As a result, at each linearization step, the following linear integral equation relates the differential optical measurements to a small perturbation α on the absorption coefficient μ_a:

$$\begin{aligned} \Gamma_{i,j} &= -\int_\Omega \overline{g_i^*(\mathbf{x})} g_j(\mathbf{x}) \alpha(\mathbf{x}) d\mathbf{x} && (2.9) \\ &:= \int_\Omega H_{i,j}(\mathbf{x}) \alpha(\mathbf{x}) d\mathbf{x} \\ &:= (\mathcal{A}_a \alpha)_{i,j}, && (2.10) \end{aligned}$$

where $H_{i,j} = -\overline{g_i^*} g_j$ is the kernel of the integral operator $\mathcal{A}_a : L^\infty(\Omega) \rightarrow \mathbb{C}^{N_d \times N_s}$, g_j is the solution of (5.6), g_i^* is the solution of (5.7), and $\Gamma_{i,j}$ is the (i,j)th entry in $\Gamma \in \mathbb{C}^{N_d \times N_s}$ which represents differential measurement at the i^{th} detector due to the j^{th} source. Note that approximating Q_i^* in (5.4) by a Gaussian function centered at $\mathbf{x}_d^i$ implies that $\Gamma_{i,j}$ corresponds to the scattered optical field evaluated at $\mathbf{x}_d^i$, after filtering it by that Gaussian function. Thus, the Gaussian approximation of the adjoint source models the finite size of the detectors. Similarly, approximating Q_j in (5.1) by a Gaussian function models the finite beam of the source.

The linear operator $\mathcal{A}_a : L^\infty(\Omega) \rightarrow \mathbb{C}^{N_d \times N_s}$ defined by (5.9) is compact and bounded by (see A.1 and A.2)

$$\|\mathcal{A}_a\|_{L^\infty(\Omega) \rightarrow l^1} \leq N_d N_s \max_i \|g_i^*\|_0 \max_j \|g_j\|_0. \tag{2.11}$$

For the given solution space $L^\infty(\Omega)$ for α, the compactness of the linear operator $\mathcal{A}_a$ implies the ill-posedness of (5.9). Hence, we regularize (5.9) with a zeroth order

Tikhonov regularization. This yields the following equation which defines our inverse problem at each linearization step:

$$\begin{aligned} \gamma = \mathcal{A}_a^* \Gamma_{i,j} &= (\mathcal{A}_a^* \mathcal{A}_a + \lambda \mathcal{I})\alpha^\lambda && (2.12) \\ &:= \mathcal{K}\alpha^\lambda, && (2.13) \end{aligned}$$

where $\lambda > 0$ and α^λ is an approximation to α. In this representation, $\mathcal{A}_a^* : \mathbb{C}^{N_d \times N_s} \to L^1(\Omega)$ is the adjoint of $\mathcal{A}_a$ and $\mathcal{I}$ is the identity operator.

Let $\mathcal{A} = \mathcal{A}_a^* \mathcal{A}_a$, then $\mathcal{A} : L^\infty(\Omega) \to L^1(\Omega)$ is defined as follows:

$$\begin{aligned} (\mathcal{A}\alpha)(\mathbf{x}) &= \sum_{i,j}^{N_d,N_s} H_{i,j}^*(\mathbf{x}) \int_\Omega H_{i,j}(\acute{\mathbf{x}})\alpha(\acute{\mathbf{x}})d\acute{\mathbf{x}} \\ &:= \int_\Omega \kappa(\mathbf{x}, \acute{\mathbf{x}})\alpha(\acute{\mathbf{x}})d\acute{\mathbf{x}}, && (2.14) \end{aligned}$$

where $\kappa(\mathbf{x}, \acute{\mathbf{x}})$ stands for the kernel of the integral operator $\mathcal{A}$ and is given by

$$\kappa(\mathbf{x}, \acute{\mathbf{x}}) = \sum_{i,j}^{N_d,N_s} H_{i,j}^*(\mathbf{x}) H_{i,j}(\acute{\mathbf{x}}), \tag{2.15}$$

and $H_{i,j}^*$ is the kernel of the adjoint operator $\mathcal{A}_a^*$ given by:

$$(\mathcal{A}_a^* \beta)(\mathbf{x}) = \sum_{i,j}^{N_d,N_s} H_{i,j}^*(\mathbf{x})\beta_{i,j} = \sum_{i,j}^{N_d,N_s} -g_i^*(\mathbf{x})\overline{g_j(\mathbf{x})}\beta_{i,j}, \tag{2.16}$$

for all $\beta \in \mathbb{C}^{N_d \times N_s}$.

Having defined the adjoint integral operator, we note that the operator $\mathcal{A} : L^\infty(\Omega) \to L^1(\Omega)$ is compact and that the operator $\mathcal{K} : L^\infty(\Omega) \to L^1(\Omega)$ is bounded by $\|\mathcal{K}\| \leq \|\mathcal{A}_a\|^2 + \lambda$. We assume that the solution $\alpha^\lambda \in L^\infty(\Omega)$ also satisfies $\alpha^\lambda \in H^1(\Omega)$. For the rest of the paper, we will denote $L^\infty(\Omega)$ and $L^1(\Omega)$ by X and Y, respectively.

2.3 Discretization of the Inverse and Forward Problems

In this section, we outline the discretization of the inverse and forward problems.

2.3.1 Inverse Problem Discretization

In practice, we seek a finite dimensional approximation to the solution of the inverse problem (5.19) at each linearization step. Therefore, we discretize (5.19) by projecting it onto a finite dimensional subspace.

Let $X_n \subset X$ and $Y_n \subset Y$ denote a sequence of finite dimensional subspaces of dimension $n = 1, 2, \ldots$, spanned by first order Lagrange basis functions $\{L_1, \ldots, L_n\}$, and $\{\mathbf{x}_p\}$, $p = 1, \ldots, n$, be the set of collocation points on Ω. Then, the collocation method approximates the solution of (5.19) by an element $\alpha_n^\lambda \in X_n$ which satisfies

$$(\mathcal{K}\alpha_n^\lambda)(\mathbf{x}_p) = \gamma(\mathbf{x}_p), \quad p = 1, \ldots, n, \tag{2.17}$$

where we express α_n^λ as

$$\alpha_n^\lambda(\mathbf{x}) = \sum_{k=1}^{n} a_k L_k(\mathbf{x}). \tag{2.18}$$

Note that in (3.10), $a_p = \alpha_n^\lambda(\mathbf{x}_p)$, $p = 1, \ldots, n$. Then, (2.17) can be written explicitly as follows:

$$\lambda a_p + \sum_{k=1}^{n} a_k \int_\Omega \kappa(\mathbf{x}_p, \acute{\mathbf{x}}) L_k(\acute{\mathbf{x}}) d\acute{\mathbf{x}} = \gamma(\mathbf{x}_p), \quad p = 1, \ldots, n. \tag{2.19}$$

Equivalently, the collocation method can be interpreted as a projection with the interpolation operator $\mathcal{P}_n : Y \to Y_n$ defined by [57]

$$\mathcal{P}_n f(\mathbf{x}) := \sum_{p=1}^{n} f(\mathbf{x}_p) L_p(\mathbf{x}), \quad \mathbf{x} \in \Omega, \tag{2.20}$$

for all $f \in Y$. Then, (2.17) is equivalent to

$$\mathcal{P}_n \mathcal{K} \alpha_n^\lambda = \mathcal{P}_n \gamma. \tag{2.21}$$

2.3.2 Forward Problem Discretization

In this section, we consider the finite element discretization of (5.6) and (5.7), and use their solutions to approximate $H_{i,j}$ and $H^*_{i,j}$. As a result, we obtain finite dimensional approximations to $\mathcal{K}$ and γ.

Let L_k be the first order Lagrange basis functions. Replacing ϕ and g_j in (5.6) with their finite dimensional counterparts:

$$\begin{aligned}\Phi(\mathbf{x}) &= \sum_{k=1}^{N_j} p_k L_k(\mathbf{x}),\\ G_j(\mathbf{x}) &= \sum_{k=1}^{N_j} c_k L_k(\mathbf{x}),\end{aligned}$$

and replacing ϕ and g^*_i in (5.7) with

$$\begin{aligned}\Phi(\mathbf{x}) &= \sum_{k=1}^{N_i} p_k L_k(\mathbf{x}),\\ G^*_i(\mathbf{x}) &= \sum_{k=1}^{N_i} d_k L_k(\mathbf{x}),\end{aligned}$$

yields the matrix equations:

$$\mathbf{S}\mathbf{c}_j = \mathbf{q}_j, \tag{2.22}$$

$$\mathbf{S}^*\mathbf{d}_i = \mathbf{q}^*_i, \tag{2.23}$$

for $\mathbf{c}_j = [c_1, c_2, \cdots, c_{N_j}]^T$ and $\mathbf{d}_i = [d_1, d_2, \cdots, d_{N_i}]^T$. Here $\mathbf{S}$ and $\mathbf{S}^*$ are the finite element matrices and $\mathbf{q}_j$ and $\mathbf{q}^*_i$ are the load vectors resulting from the finite element discretization of (5.6) and (5.7). Note that for each source (detector), the dimension of the finite element solution G_j (G^*_i) can be different, therefore N_j (N_i) may vary.

The $H^1(\Omega)$ boundedness of the solutions g_j and g^*_i implies that the discretization error e_j and e^*_i in G_j and G^*_i on the m^{th} finite element is bounded. A bound for e_j and e^*_i on each finite element can be found by using the discretization error estimates (Theorem 4.4.4 in [20, 9]):

$$\|e_j\|_{0,m} \leq C\|g_j\|_{1,m} h_m^j, \tag{2.24}$$

$$\|e_i^*\|_{0,m} \leq C\|g_i^*\|_{1,m} h_m^i, \tag{2.25}$$

where C is a positive constant and h_m^j (h_m^i) is the diameter of the smallest ball containing the m^{th} finite element.

2.3.3 Discretization of the Inverse Problem with Operator Approximations

Substituting the finite element approximations G_j and G_i^* in (4.23) and (5.15), and using the resulting finite dimensional operator approximations in (3.9), we obtain the following linear system in terms of $\tilde{\alpha}_n^\lambda$ which approximates α^λ:

$$\mathcal{P}_n \tilde{\mathcal{K}} \tilde{\alpha}_n^\lambda = \mathcal{P}_n \tilde{\gamma}. \tag{2.26}$$

In (2.26), the operator $\tilde{\mathcal{K}} : X \rightarrow Y$ is the finite dimensional approximation of $\mathcal{K}$ in (5.19) and $\mathcal{P}_n \tilde{\mathcal{K}} : X_n \rightarrow Y_n$. Similarly,

$$\tilde{\gamma} = \tilde{\mathcal{A}}_a^* \Gamma, \tag{2.27}$$

where $\tilde{\mathcal{A}}_a^*$ is the approximation to the adjoint operator $\mathcal{A}_a^*$, obtained by substituting G_j and G_i^* in (5.15).

2.4 Discretization-based Error Analysis

As a result of the discretization of the forward and inverse problems, the reconstructed image $\tilde{\alpha}_n^\lambda$ in (2.26) is an approximation to the actual image α^λ. Thus, the accuracy of the reconstructed image depends on the error incurred by the discretization of the forward and inverse problems.

In this section, we analyze the effect of the discretization of the forward and inverse problems on the accuracy of DOT imaging. The analysis is carried out based on the inverse problem at each linearization defined by (5.19) and the associated kernel $\kappa(\mathbf{x}, \acute{\mathbf{x}})$.

In this work, we follow an approach which allows us to separately analyze the effect of the discretization of each problem on the accuracy of the reconstructed optical image. In this respect, we first consider the impact of projection (i.e. inverse problem discretization) by collocation method when the associated kernel $\kappa(\mathbf{x}, \acute{\mathbf{x}})$ in (5.19) is exact. Next, we explore the case in which the kernel is replaced by its finite dimensional approximation (i.e. degenerate kernel) and analyze the effect of the forward problem discretization on the accuracy of the reconstructed image without projecting (5.19).

Our analysis reveals that even the kernel is exact, the accuracy of the solution approximation α_n^λ in (3.9) resulting from the inverse problem discretization depends on the kernel $\kappa(\mathbf{x}, \acute{\mathbf{x}})$ of the integral operator. Likewise, the error in the reconstructed optical image due to the discretization of the forward problem is a function of the inverse problem solution. These results suggest that the discretization of the inverse and forward problems can not be considered independent of each other.

2.4.1 Case 1: The kernel $\kappa(\mathbf{x}, \acute{\mathbf{x}})$ is exact

In this section, we show the effect of projection on the optical imaging accuracy. In the analysis, we assume that the kernel $\kappa(\mathbf{x}, \acute{\mathbf{x}})$ is exact. We first prove the convergence of the projection method for the operator $\mathcal{K}$, and then analyze the effect of projection on the imaging accuracy.

Clearly, the inverse operator $\mathcal{K}^{-1} : Y \to X$ exists since $\mathcal{K}$ is positive definite for $\lambda > 0$. Furthermore, by the compactness of $\mathcal{A}$ and Riesz Theorem, the inverse operator $\mathcal{K}^{-1}$ is bounded by

$$\|\mathcal{K}^{-1}\| \leq \frac{1}{\lambda}. \tag{2.28}$$

Lemma: Projection by collocation method for the operator $\mathcal{K} : X \to Y$ converges. Specifically, the sequence of finite dimensional operators $\mathcal{P}_n\mathcal{K} : X_n \to Y_n$ are invertible for sufficiently large n, and $(\mathcal{P}_n\mathcal{K})^{-1}\mathcal{P}_n\mathcal{K}\alpha^\lambda \to \alpha^\lambda$, $n \to \infty$. Furthermore,

$$\|(\mathcal{P}_n\mathcal{K})^{-1}\mathcal{P}_n\mathcal{K}\|_{X \to X_n} \leq C_M \frac{\|\mathcal{K}\|_{X \to Y}}{\lambda} \tag{2.29}$$

for some $C_M > 0$ independent of n.

Proof. See A.3. ☐

Based on the Lemma, the following theorem provides an upper bound for the $L^1(\Omega)$ norm of the error between the solution α^λ of (5.19) and the solution α_n^λ of (3.9).

Theorem 1: Let $\{\Omega_m\}$ denote a set of linear finite elements for $m = 1, \cdots, N_\Delta$, such that $\bigcup_m^{N_\Delta} \Omega_m = \Omega$ and h_m be the diameter of the smallest ball that contains the m^{th} element. Then,

$$\begin{aligned} \|\alpha^\lambda - \alpha_n^\lambda\|_{L^1(\Omega)} &\leq C\sqrt{V_\Omega}\|\mathcal{I} - \mathcal{T}_n\|_{Y\to X_n} \sum_{m=1}^{N_\Delta} \|\alpha^\lambda\|_{1,m} h_m \\ &+ \frac{C}{\lambda}\|\mathcal{T}_n\|_{Y\to X_n} \max_{i,j} \|g_i^* g_j\|_{L^1(\Omega)} \sum_{m=1}^{N_\Delta} \sum_{i,j}^{N_d,N_s} \|g_i^* g_j\|_{0,m} \|\alpha^\lambda\|_{1,m} h_m, \end{aligned} \tag{2.30}$$

where C is a positive constant, V_Ω is the volume of Ω and $\mathcal{T}_n : Y \to X_n$ is a uniformly bounded operator given by $\mathcal{T}_n = (\mathcal{I} + \frac{1}{\lambda}\mathcal{P}_n\mathcal{A})^{-1}\mathcal{P}_n$.

Proof.

$$\begin{aligned} \alpha^\lambda - \alpha_n^\lambda &= [\mathcal{I} - (\mathcal{P}_n\mathcal{K})^{-1}\mathcal{P}_n\mathcal{K}]\alpha^\lambda \\ &= [\mathcal{I} - (\mathcal{P}_n\mathcal{K})^{-1}\mathcal{P}_n\mathcal{K}](\alpha^\lambda - \psi) \end{aligned} \tag{2.31}$$

since $[\mathcal{I} - (\mathcal{P}_n\mathcal{K})^{-1}\mathcal{P}_n\mathcal{K}]\psi = 0$, where $\psi \in X_n$ is the interpolant of α^λ [20, 9]. Using (A.6),

$$\begin{aligned} [\mathcal{I} - (\mathcal{P}_n\mathcal{K})^{-1}\mathcal{P}_n\mathcal{K}] &= \mathcal{I} - (\mathcal{I} + \frac{1}{\lambda}\mathcal{P}_n\mathcal{A})^{-1}\frac{1}{\lambda}\mathcal{P}_n\mathcal{K} \\ &= \mathcal{I} - \mathcal{T}_n\frac{1}{\lambda}\mathcal{K}, \end{aligned} \tag{2.32}$$

where $\mathcal{T}_n = (\mathcal{I} + \frac{1}{\lambda}\mathcal{P}_n\mathcal{A})^{-1}\mathcal{P}_n$ is a uniformly bounded operator (see A.3). We use $\mathcal{K}$ defined by (5.19) and (2.32) in (2.31) to obtain

$$\alpha^\lambda - \alpha_n^\lambda = (\mathcal{I} - \mathcal{T}_n)(\alpha^\lambda - \psi) - \frac{\mathcal{T}_n}{\lambda}\mathcal{A}(\alpha^\lambda - \psi). \tag{2.33}$$

Then we use the definition of $\mathcal{A}$ in (2.33) and find

$$\alpha^\lambda - \alpha_n^\lambda = (\mathcal{I} - \mathcal{T}_n)(\alpha^\lambda - \psi) - \frac{\mathcal{T}_n}{\lambda} \int_\Omega \kappa(\cdot, \acute{\mathbf{x}})(\alpha^\lambda - \psi)(\acute{\mathbf{x}}) d\acute{\mathbf{x}}. \tag{2.34}$$

This leads to

$$\begin{aligned} \|\alpha^\lambda - \alpha_n^\lambda\|_{L^1(\Omega)} &\leq \|\mathcal{I} - \mathcal{T}_n\|_{Y \to X_n} \|\alpha^\lambda - \psi\|_{L^1(\Omega)} \\ &+ \frac{1}{\lambda} \|\mathcal{T}_n\|_{Y \to X_n} \| \int_\Omega \kappa(\cdot, \acute{\mathbf{x}})(\alpha^\lambda - \psi)(\acute{\mathbf{x}}) d\acute{\mathbf{x}} \|_{L^1(\Omega)} \\ &\leq \sqrt{V_\Omega} \|\mathcal{I} - \mathcal{T}_n\|_{Y \to X_n} \|\alpha^\lambda - \psi\|_0 \\ &+ \frac{1}{\lambda} \|\mathcal{T}_n\|_{Y \to X_n} \int_\Omega d\mathbf{x} \int_\Omega |\kappa(\mathbf{x}, \acute{\mathbf{x}})(\alpha^\lambda - \psi)(\acute{\mathbf{x}})| d\acute{\mathbf{x}}, \end{aligned} \tag{2.35}$$

The second term in (2.35) can be rewritten as:

$$\begin{aligned} &\frac{1}{\lambda} \|\mathcal{T}_n\|_{Y \to X_n} \int_\Omega d\mathbf{x} \int_\Omega |\kappa(\mathbf{x}, \acute{\mathbf{x}})(\alpha^\lambda - \psi)(\acute{\mathbf{x}})| d\acute{\mathbf{x}} \\ = &\frac{1}{\lambda} \|\mathcal{T}_n\|_{Y \to X_n} \int_\Omega d\mathbf{x} \left(\sum_{m=1}^{N_\Delta} \int_{\Omega_m} |\kappa(\mathbf{x}, \acute{\mathbf{x}})(\alpha^\lambda - \psi)(\acute{\mathbf{x}})| d\acute{\mathbf{x}} \right). \end{aligned} \tag{2.36}$$

Let e_α be the interpolation error:

$$e_\alpha = \alpha^\lambda - \psi. \tag{2.37}$$

Then, using (4.18),

$$\begin{aligned} &\sum_{m=1}^{N_\Delta} \int_{\Omega_m} |\kappa(\mathbf{x}, \acute{\mathbf{x}}) e_\alpha(\acute{\mathbf{x}})| d\acute{\mathbf{x}} \\ = &\sum_{m=1}^{N_\Delta} \int_{\Omega_m} \left| \sum_{i,j}^{N_d, N_s} g_i^*(\mathbf{x}) \overline{g_j(\mathbf{x}) g_i^*(\acute{\mathbf{x}})} g_j(\acute{\mathbf{x}}) e_\alpha(\acute{\mathbf{x}}) \right| d\acute{\mathbf{x}} \end{aligned} \tag{2.38}$$

$$\begin{aligned} &\leq \sum_{m=1}^{N_\Delta} \sum_{i,j}^{N_d, N_s} |g_i^*(\mathbf{x}) g_j(\mathbf{x})| \int_{\Omega_m} |g_i^*(\acute{\mathbf{x}}) g_j(\acute{\mathbf{x}})| |e_\alpha(\acute{\mathbf{x}})| d\acute{\mathbf{x}} \\ &\leq \sum_{m=1}^{N_\Delta} \sum_{i,j}^{N_d, N_s} |g_i^*(\mathbf{x}) g_j(\mathbf{x})| \|g_i^* g_j\|_{0,m} \|e_\alpha\|_{0,m}, \end{aligned} \tag{2.39}$$

where (2.39) follows from the Schwarz' inequality. Note that $g_i^* g_j \in L^2(\Omega)$ by assuming (5.8) holds up to the boundary $\partial\Omega$.

We now use (2.36) and (2.39) to obtain

$$\frac{1}{\lambda}\|\mathcal{T}_n\|_{Y\to X_n}\int_\Omega d\mathbf{x}\left(\ \textstyle\int_\Omega |\kappa(\mathbf{x},\acute{\mathbf{x}})(\alpha^\lambda-\psi)(\acute{\mathbf{x}})|d\acute{\mathbf{x}}\ \right) \leq \frac{1}{\lambda}\|\mathcal{T}_n\|_{Y\to X_n} \times \int_\Omega d\mathbf{x}\sum_{m=1}^{N_\Delta}\sum_{i,j}^{N_d,N_s}|g_i^*(\mathbf{x})g_j(\mathbf{x})|\|g_i^*g_j\|_{0,m}\|e_\alpha\|_{0,m}. \tag{2.40}$$

Using the bound (2.40) in (2.35) and substituting the interpolation error bound [20, 9]

$$\|e_\alpha\|_{0,m} \leq C\|\alpha^\lambda\|_{1,m}h_m, \tag{2.41}$$

we obtain

$$\begin{aligned}\|\alpha^\lambda-\alpha_n^\lambda\|_{L^1(\Omega)} &\leq C\sqrt{V_\Omega}\|\mathcal{I}-\mathcal{T}_n\|_{Y\to X_n}\sum_{m=1}^{N_\Delta}\|\alpha^\lambda\|_{1,m}h_m \\ &+\frac{C}{\lambda}\|\mathcal{T}_n\|_{Y\to X_n}\sum_{m=1}^{N_\Delta}\sum_{i,j}^{N_d,N_s}\|g_i^*g_j\|_{L^1(\Omega)}\|g_i^*g_j\|_{0,m}\|\alpha^\lambda\|_{1,m}h_m. \\ &\leq C\sqrt{V_\Omega}\|\mathcal{I}-\mathcal{T}_n\|_{Y\to X_n}\sum_{m=1}^{N_\Delta}\|\alpha^\lambda\|_{1,m}h_m \\ &+\frac{C}{\lambda}\|\mathcal{T}_n\|_{Y\to X_n}\max_{i,j}\|g_i^*g_j\|_{L^1(\Omega)}\sum_{m=1}^{N_\Delta}\sum_{i,j}^{N_d,N_s}\|g_i^*g_j\|_{0,m}\|\alpha^\lambda\|_{1,m}h_m.\end{aligned} \tag{2.42}$$

□

Remarks 1:

1. Theorem 1 shows the spatial dependence of the inverse problem discretization on the forward problem solution.
2. The first term in (2.42) suggests that the mesh of the inverse problem be refined where $\|\alpha^\lambda\|_1$ is large.

3. The second term in (2.42) shows that the term $\|\alpha^\lambda\|_{1,m}$ is scaled spatially by $\|g_i^* g_j\|_{0,m}$. Thus, the effect of the interpolation error e_α (see equation (3.11)) in the inverse problem solution is scaled spatially by the solution of the forward problem. As a result, the orientation of the sources and detectors with respect to the support of the optical heterogeneity determines the extent of the bound on $\|\alpha^\lambda - \alpha_n^\lambda\|_{L^1(\Omega)}$.
4. The regularization parameter affects the bound on $\|\alpha^\lambda - \alpha_n^\lambda\|_{L^1(\Omega)}$.
5. Increasing the number of sources and detectors increases the bound on $\|\alpha^\lambda - \alpha_n^\lambda\|_{L^1(\Omega)}$.

Remarks 2:

1. Note that the conventional interpolation error estimate given in (2.41) depends on only the smoothness and support of α^λ, and the finite dimensional space of approximating functions [20]. On the other hand, the error estimate (3.12) in Theorem 1 shows that the accuracy of the reconstructed image α_n^λ depends on the orientation of the absorptive heterogeneity with respect to the sources and detectors, as well as on the bound (2.41) on the interpolation error.
2. An error bound similar to (3.12) follows if one uses the Galerkin method [57] instead of the collocation method for projection.
3. The interpolation error bound (2.41) can be modified based on the choice of the basis function in (3.10) and the smoothness of the solution α^λ (Theorem 4.4.4. in [20, 9]). For instance, if $\alpha^\lambda \in H^2(\Omega)$ and quadratic Lagrange basis functions are used, then (2.41) can be replaced by

$$\|e_\alpha\|_{0,m} \leq C\|\alpha^\lambda\|_{2,m} h_m^2.$$

4. An error bound similar to (3.12) can be derived for the error that occurs as a result of the discretization of the inverse problem in electrical impedance tomography, optical fluorescence tomography, bioluminescence tomography, and microwave imaging. Note that in all these imaging modalities, the forward

problem is modeled by a PDE and the inverse problem can be interpreted in terms of a linear integral equation, whose kernel is related to the solution of this PDE.

5. Let γ^δ be the perturbed right-hand side γ of (3.9) due to the presence of noise, such that $\|\gamma^\delta - \gamma\|_{L^1(\Omega)} \leq \delta$. Then, an additional term is introduced to the error bound in (3.12) due to this perturbation:

$$\|\alpha^\lambda - \alpha_n^\lambda\|_{L^1(\Omega)} \leq C\sqrt{V_\Omega}\|\mathcal{I} - \mathcal{T}_n\|_{Y\to X_n} \sum_{m=1}^{N_\Delta} \|\alpha^\lambda\|_{1,m} h_m$$
$$+\frac{C}{\lambda}\|\mathcal{T}_n\|_{Y\to X_n} \max_{i,j} \|g_i^* g_j\|_{L^1(\Omega)} \sum_{m=1}^{N_\Delta} \sum_{i,j}^{N_d,N_s} \|g_i^* g_j\|_{0,m} \|\alpha^\lambda\|_{1,m} h_m$$
$$+\frac{C_M}{\lambda}\delta, \tag{2.43}$$

where $C_M > 0$ is the constant in (2.29) with the use of first order Lagrange basis functions (see A.3). Thus, the additional term $C_M\delta/\lambda$ indicates that the choice of basis functions may be critical in the presence of noise.

2.4.2 Case 2: The kernel is degenerate

In this section, we first derive approximate upper bounds for the approximation errors $\|\tilde{\mathcal{K}} - \mathcal{K}\|$ and $\|\tilde{\gamma} - \gamma\|$, which result from the discretization of the forward problem. Then, we show the effect of these approximation errors on the accuracy of the reconstructed optical image. For notational convenience, we will drop the subscripts on the norms $\|\cdot\|$ where necessary.

The operator $\mathcal{K} : X \to Y$ is bounded with a bounded inverse $\mathcal{K}^{-1} : Y \to X$. By the finite element approximation of the associated kernel, the sequence of bounded linear finite dimensional operators $\tilde{\mathcal{K}}$ is norm convergent $\|\tilde{\mathcal{K}} - \mathcal{K}\| \to 0; N_j, N_i \to \infty$, for $j = 1, \cdots, N_s$ and $i = 1, \cdots, N_d$, and

$$\|\tilde{\mathcal{K}}^{-1}\| < 1/\lambda, \tag{2.44}$$

which can be obtained analogous to (4.33).

In the following, we derive an explicit approximation to the error $\|\tilde{\mathcal{K}} - \mathcal{K}\|$

in terms of the associated kernel and the discretization error in the kernel approximation. The result is then used to compute the error in the reconstructed optical image due to $\|\tilde{\mathcal{K}} - \mathcal{K}\|$.

By definition,

$$\|(\mathcal{A}_a - \tilde{\mathcal{A}}_a)\alpha\|_{l^1} = \sum_{i,j}^{N_d, N_s} \left| \int_\Omega (\overline{g_i^*(\mathbf{x})} g_j(\mathbf{x}) - \overline{G_i^*(\mathbf{x})} G_j(\mathbf{x}))\alpha(\mathbf{x}) d\mathbf{x} \right|, \tag{2.45}$$

where G_i^*, G_j are finite element approximations to g_i^* and g_j, respectively. We can expand $\overline{g_i^*} g_j - \overline{G_i^*} G_j$ as

$$\overline{g_i^*} g_j - \overline{G_i^*} G_j = \overline{e_i^*} e_j + G_j \overline{e_i^*} + \overline{G_i^*} e_j, \tag{2.46}$$

where $e_i^* = g_i^* - G_i^*$ and $e_j = g_j - G_j$. Replacing G_i^* and G_j respectively with $g_i^* - e_i^*$ and $g_j - e_j$, we get

$$\begin{aligned} \overline{g_i^*} g_j - \overline{G_i^*} G_j &= g_j \overline{e_i^*} + \overline{g_i^*} e_j - \overline{e_i^*} e_j \\ &\approx g_j \overline{e_i^*} + \overline{g_i^*} e_j, \end{aligned} \tag{2.47}$$

where we neglect the term $\overline{e_i^*} e_j$.

We can express $\mathcal{K} - \tilde{\mathcal{K}}$ as

$$\mathcal{K} - \tilde{\mathcal{K}} = \mathcal{A}_a^* \mathcal{A}_a - \tilde{\mathcal{A}}_a^* \tilde{\mathcal{A}}_a. \tag{2.48}$$

Following a similar approach as above,

$$\mathcal{A}_a^* \mathcal{A}_a - \tilde{\mathcal{A}}_a^* \tilde{\mathcal{A}}_a = (\mathcal{A}_a^* - \tilde{\mathcal{A}}_a^*)(\mathcal{A}_a - \tilde{\mathcal{A}}_a) + \tilde{\mathcal{A}}_a^*(\mathcal{A}_a - \tilde{\mathcal{A}}_a) + (\mathcal{A}_a^* - \tilde{\mathcal{A}}_a^*)\tilde{\mathcal{A}}_a. \tag{2.49}$$

As a result, the following condition holds:

$$\|\tilde{\mathcal{K}} - \mathcal{K}\| \leq \|(\mathcal{A}_a^* - \tilde{\mathcal{A}}_a^*)(\mathcal{A}_a - \tilde{\mathcal{A}}_a)\| + \|\tilde{\mathcal{A}}_a^*(\mathcal{A}_a - \tilde{\mathcal{A}}_a) + (\mathcal{A}_a^* - \tilde{\mathcal{A}}_a^*)\tilde{\mathcal{A}}_a\|. \tag{2.50}$$

Since $\tilde{\mathcal{A}}_a = -(\mathcal{A}_a - \tilde{\mathcal{A}}_a) + \mathcal{A}_a$, (C.9) can be rewritten as

$$\begin{aligned}
\|\tilde{\mathcal{K}} - \mathcal{K}\| &= \|\mathcal{A}_a^*\mathcal{A}_a - \tilde{\mathcal{A}}_a^*\tilde{\mathcal{A}}_a\| \\
&\leq \|(\mathcal{A}_a^* - \tilde{\mathcal{A}}_a^*)(\mathcal{A}_a - \tilde{\mathcal{A}}_a)\| + 2\|\mathcal{A}_a^*(\mathcal{A}_a - \tilde{\mathcal{A}}_a)\| \\
&\approx 2\|\mathcal{A}_a^*(\mathcal{A}_a - \tilde{\mathcal{A}}_a)\|, \qquad (2.51)
\end{aligned}$$

where we neglect the term $\|(\mathcal{A}_a^* - \tilde{\mathcal{A}}_a^*)(\mathcal{A}_a - \tilde{\mathcal{A}}_a)\|$.

Similarly, $\|\tilde{\gamma} - \gamma\|$ can be interpreted as

$$\begin{aligned}
\|\tilde{\gamma} - \gamma\|_{L^1(\Omega)} &= \int_\Omega \left| \sum_{i,j}^{N_d,N_s} (g_i^*(\mathbf{x})\overline{g_j(\mathbf{x})} - G_i^*(\mathbf{x})\overline{G_j(\mathbf{x})})\Gamma_{i,j} \right| d\mathbf{x} \\
&\approx \int_\Omega \left| \sum_{i,j}^{N_d,N_s} (e_i^*(\mathbf{x})\overline{g_j(\mathbf{x})} + g_i^*(\mathbf{x})\overline{e_j(\mathbf{x})})\Gamma_{i,j} \right| d\mathbf{x}, \qquad (2.52)
\end{aligned}$$

where the error in $\Gamma_{i,j}$ due to discretization is neglected and the last approximation is derived similar to (C.10).

We now analyze the effect of the forward problem discretization on the accuracy of the reconstructed optical image. Let $\tilde{\alpha}^\lambda$ be the solution of

$$\tilde{\mathcal{K}}\tilde{\alpha}^\lambda = \tilde{\gamma}, \qquad (2.53)$$

where $\tilde{\mathcal{K}}$ and $\tilde{\gamma}$ are the finite dimensional approximations to $\mathcal{K}$ and γ, respectively. Then, by Theorem 10.1 in [57], the error in the solution $\tilde{\alpha}^\lambda$ with respect to the actual solution α^λ is bounded by

$$\|\alpha^\lambda - \tilde{\alpha}^\lambda\| \leq \frac{1}{\lambda}\left\{ \|(\tilde{\mathcal{K}} - \mathcal{K})\alpha^\lambda\| + \|\tilde{\gamma} - \gamma\| \right\}. \qquad (2.54)$$

In the next theorem, we will expand the terms in (4.32) to show explicitly the effect of the forward problem discretization on the accuracy of the inverse problem solution.

Theorem 2: Let $\{\Omega_m^j\}$ denote the set of linear elements used to discretize (5.6) for $m = 1, \ldots, N_\Delta^j$; such that $\bigcup_m^{N_\Delta^j} \Omega_m^j = \Omega$ and h_m^j be the diameter of the smallest ball that contains the m^{th} element in the solution G_j, for all $j = 1, \ldots N_s$. Similarly, let $\{\Omega_n^i\}$ denote the set of linear elements used to discretize (5.7) for $n = 1, \ldots, N_\Delta^{*i}$; such that $\bigcup_n^{N_\Delta^{*i}} \Omega_n^i = \Omega$ and h_n^i be the diameter of the smallest ball that contains the n^{th} element in the solution G_i^*, for all $i = 1, \ldots N_d$. Then, a bound for the error between the solution α^λ of (5.19) and the solution $\tilde{\alpha}^\lambda$ of (3.16) due to the approximations $\tilde{\mathcal{K}}$ and $\tilde{\gamma}$ is given by:

$$\begin{aligned}
\|\alpha^\lambda - \tilde{\alpha}^\lambda\|_{L^1(\Omega)} \leq\ & \frac{C}{\lambda} \max_{i,j} \|g_i^* g_j\|_{L^1(\Omega)} \\
\times\ & \left(\sum_{i=1}^{N_d} \sum_{n,j}^{N_\Delta^{*i}, N_s} \left(2\|g_j \alpha^\lambda\|_{0,n^i} + \|\alpha\|_\infty \|g_j\|_{0,n^i} \right) \|g_i^*\|_{1,n^i} h_n^i \right. \\
+\ & \left. \sum_{j=1}^{N_s} \sum_{m,i}^{N_\Delta^j, N_d} \left(2\|g_i^* \alpha^\lambda\|_{0,m^j} + \|\alpha\|_\infty \|g_i^*\|_{0,m^j} \right) \|g_j\|_{1,m^j} h_m^j \right) \\
,\ &
\end{aligned} \tag{2.55}$$

where C is a positive constant, and $\|\cdot\|_{0,a^b}$ and $\|\cdot\|_{1,a^b}$ denote respectively the L^2 and H^1 norms on the finite element Ω_a^b.

Proof. Using (2.51), (C.2), and (C.10), we can write

$$\begin{aligned}
& \|(\tilde{\mathcal{K}} - \mathcal{K})\alpha^\lambda\|_{L^1(\Omega)} \approx 2\|\mathcal{A}_a^*(\mathcal{A}_a - \tilde{\mathcal{A}}_a)\alpha^\lambda\|_{L^1(\Omega)} \\
\approx 2 & \left\| \sum_{i,j}^{N_d,N_s} g_i^*(\cdot)\overline{g_j(\cdot)} \int_\Omega \left(g_j(\acute{\mathbf{x}})\overline{e_i^*(\acute{\mathbf{x}})} + \overline{g_i^*(\acute{\mathbf{x}})} e_j(\acute{\mathbf{x}}) \right) \alpha^\lambda(\acute{\mathbf{x}}) d\acute{\mathbf{x}} \right\|_{L^1(\Omega)} \\
\leq 2 & \max_{i,j} \|g_i^* g_j\|_{L^1(\Omega)} \sum_{i,j}^{N_d,N_s} \int_\Omega \left| \left(g_j(\acute{\mathbf{x}})\overline{e_i^*(\acute{\mathbf{x}})} + \overline{g_i^*(\acute{\mathbf{x}})} e_j(\acute{\mathbf{x}}) \right) \alpha^\lambda(\acute{\mathbf{x}}) \right| d\acute{\mathbf{x}}.
\end{aligned} \tag{2.56}$$

An upper bound for the integral in (C.14) can be obtained as follows:

$$\begin{aligned}
& \int_\Omega \left| \left(g_j(\acute{\mathbf{x}})\overline{e_i^*(\acute{\mathbf{x}})} + \overline{g_i^*(\acute{\mathbf{x}})} e_j(\acute{\mathbf{x}}) \right) \alpha^\lambda(\acute{\mathbf{x}}) \right| d\acute{\mathbf{x}} \\
\leq & \sum_{n=1}^{N_\Delta^{*i}} \|e_i^*\|_{0,n^i} \|g_j \alpha^\lambda\|_{0,n^i} + \sum_{m=1}^{N_\Delta^j} \|e_j\|_{0,m^j} \|g_i^* \alpha^\lambda\|_{0,m^j}.
\end{aligned} \tag{2.57}$$

Note that $g_j \alpha^\lambda \in L^2(\Omega)$ since $|g_j \alpha^\lambda| \leq |g_j| \|\alpha^\lambda\|_\infty$. Similarly, $g_i^* \alpha^\lambda \in L^2(\Omega)$

since $|g_i^* \alpha^\lambda| \leq |g_i^*| \|\alpha^\lambda\|_\infty$. Using (C.15) in (C.14),

$$\begin{aligned}
\|(\tilde{\mathcal{K}} - \mathcal{K})\alpha^\lambda\|_{L^1(\Omega)} \leq{}& 2 \max_{i,j} \|g_i^* g_j\|_{L^1(\Omega)} \\
& \times \left(\sum_{i=1}^{N_d} \sum_{n,j}^{N_\Delta^{*i}, N_s} \|e_i^*\|_{0,n^i} \|g_j \alpha^\lambda\|_{0,n^i} \right. \\
& \left. + \sum_{j=1}^{N_s} \sum_{m,i}^{N_\Delta^j, N_d} \|e_j\|_{0,m^j} \|g_i^* \alpha^\lambda\|_{0,m^j} \right).
\end{aligned} \tag{2.58}$$

To compute an upper bound for $\|\tilde{\gamma} - \gamma\|$ using (4.38), we first write

$$\begin{aligned}
&\int_\Omega \left| \sum_{i,j}^{N_d, N_s} \left(e_i^*(\mathbf{x}) \overline{g_j(\mathbf{x})} + g_i^*(\mathbf{x}) \overline{e_j(\mathbf{x})} \right) \Gamma_{i,j} \right| d\mathbf{x} \\
&\leq \max_{i,j} |\Gamma_{i,j}| \int_\Omega \sum_{i,j}^{N_d, N_s} |e_i^*(\mathbf{x}) \overline{g_j(\mathbf{x})} + g_i^*(\mathbf{x}) \overline{e_j(\mathbf{x})}| d\mathbf{x} \\
&\leq \max_{i,j} |\Gamma_{i,j}| \left(\sum_{i=1}^{N_d} \sum_{n,j}^{N_\Delta^{*i}, N_s} \|e_i^*\|_{0,n^i} \|g_j\|_{0,n^i} + \sum_{i=j}^{N_s} \sum_{m,i}^{N_\Delta^j, N_d} \|g_i^*\|_{0,m^j} \|e_j\|_{0,m^j} \right).
\end{aligned} \tag{2.59}$$

Noting (5.9),

$$\max_{i,j} |\Gamma_{i,j}| \leq \max_{i,j} \|g_i^* g_j\|_{L^1(\Omega)} \|\alpha\|_\infty, \tag{2.60}$$

which leads to

$$\begin{aligned}
&\max_{i,j} |\Gamma_{i,j}| \left(\sum_{i=1}^{N_d} \sum_{n,j}^{N_\Delta^{*i}, N_s} \|e_i^*\|_{0,n^i} \|g_j\|_{0,n^i} + \sum_{i=j}^{N_s} \sum_{m,i}^{N_\Delta^j, N_d} \|g_i^*\|_{0,m^j} \|e_j\|_{0,m^j} \right) \\
&\leq \max_{i,j} \|g_i^* g_j\|_{L^1(\Omega)} \|\alpha\|_\infty \left(\sum_{i=1}^{N_d} \sum_{n,j}^{N_\Delta^{*i}, N_s} \|e_i^*\|_{0,n^i} \|g_j\|_{0,n^i} + \sum_{i=j}^{N_s} \sum_{m,i}^{N_\Delta^j, N_d} \|g_i^*\|_{0,m^j} \|e_j\|_{0,m^j} \right).
\end{aligned} \tag{2.61}$$

We now use (C.16), (C.24), the corresponding discretization error estimates (5.25)-

(5.26), and (4.32) to obtain (4.34). □

Remarks 3:

1. Theorem 2 suggests the use of meshes designed individually for the solutions G_j, $j = 1, \cdots, N_s$ and G_i^*, $i = 1, \cdots, N_d$.

2. Theorem 2 states explicitly the effect of the forward problem discretization on the accuracy of the inverse problem solution. In this context, Theorem 2 suggests a discretization scheme for the forward problem, where the discretization criterion is based on the inverse problem solution accuracy, rather than the accuracy of the forward problem solution.

3. For each source, when solving for G_j, h_m^j has to be kept small where $(2\|g_i^*\alpha^\lambda\|_{0,m^j} + \|\alpha\|_\infty \|g_i^*\|_{0,m^j})\|g_j\|_{1,m^j}$ is large. Note that $\|g_j\|_{1,m^j}$ will be large on the elements close to the j^{th} source.

4. For each detector, when solving for G_i^*, h_n^i has to be kept small where $(2\|g_j\alpha^\lambda\|_{0,n^i} + \|\alpha\|_\infty \|g_j\|_{0,n^i})\|g_i^*\|_{1,n^i}$ is large. Note that $\|g_i^*\|_{1,n^i}$ will be large on the elements close to the i^{th} detector.

5. $|g_j|$ and $|g_i^*|$ are higher close to the sources and detectors, respectively. Therefore, h_m^j has to be small around the j^{th} source and around all detectors, where α^λ is nonzero. Likewise, h_m^i has to be small around the i^{th} detector and around all sources, where α^λ is nonzero.

6. If α^λ is nonzero on the whole domain Ω, then the error may become higher depending on the magnitude of $|g_j|$ and $|g_i^*|$.

7. The regularization parameter affects the bound on $\|\alpha^\lambda - \tilde{\alpha}^\lambda\|_{L^1(\Omega)}$.

8. Increasing the number of sources and detectors increases the bound on $\|\alpha^\lambda - \tilde{\alpha}^\lambda\|_{L^1(\Omega)}$.

Remarks 4:

1. Note that the finite-element discretization error estimates (5.25)-(5.26) depend on only the smoothness and support of g_j and g_i^*, and the finite dimensional space of approximating functions [20]. However, the error estimate (4.34) in Theorem 2 shows that the accuracy of the reconstructed image $\tilde{\alpha}^\lambda$ depends on the orientation of the absorptive heterogeneity with respect to the sources and detectors, as well as on the finite-element discretization error estimates (5.25)-(5.26). In this respect, the estimate (4.34) in Theorem 2 shows that reducing the discretization error in the solutions G_j and G_i^* of the forward problem may not ensure the accuracy of the reconstructed absorption image (see [40]).

2. In case a different discretization approach such as finite difference [78] or finite volume [31] is used to solve the forward problem, Theorem 2 can be modified in a straightforward manner by replacing the discretization error estimates (5.25) and (5.26) with the corresponding error estimates specific to the method of choice [78], [31].

3. Let $\tilde{\gamma}^\delta$ be the perturbed right-hand side $\tilde{\gamma}$ of (3.16) due to the presence of noise, such that $\|\tilde{\gamma}^\delta - \tilde{\gamma}\|_{L^1(\Omega)} \leq \tilde{\delta}$. Then, an additional term is introduced to the bound in (4.34) due to this perturbation:

$$\begin{aligned} \|\alpha^\lambda - \tilde{\alpha}^\lambda\|_{L^1(\Omega)} \leq \frac{C}{\lambda} \max_{i,j} \|g_i^* g_j\|_{L^1(\Omega)} \\ \times \left(\sum_{i=1}^{N_d} \sum_{n,j}^{N_\Delta^{*i}, N_s} \Big(2\|g_j \alpha^\lambda\|_{0,n^i} + \|\alpha\|_\infty \|g_j\|_{0,n^i} \Big) \|g_i^*\|_{1,n^i} h_n^i \right. \\ \left. + \sum_{j=1}^{N_s} \sum_{m,i}^{N_\Delta^j, N_d} \Big(2\|g_i^* \alpha^\lambda\|_{0,m^j} + \|\alpha\|_\infty \|g_i^*\|_{0,m^j} \Big) \|g_j\|_{1,m^j} h_m^j \right) + \frac{\tilde{\delta}}{\lambda}. \end{aligned} \tag{2.62}$$

   Clearly, the additional term $\tilde{\delta}/\lambda$ due to the presence of noise in (2.62) is independent of the discretization of the forward problem.

4. Theorem 2 provides a general framework to analyze the error in reconstructed optical images resulting from the perturbations in the kernel of the linear integral equation (4.18). In general, a perturbation in the kernel of the linear integral equation (4.18) can occur due to errors resulting from the numerical

integration of (5.6)-(5.7), the approximation of the boundary $\partial\Omega$, the inaccurate approximation of the source Q_j and/or the background optical properties. Furthermore, the analysis framework in Theorem 2 can be used to analyze the effect of linearization of the Lippmann-Schwinger type equations [25] using Born approximation on the accuracy of the reconstructed optical images [38].

5. A bound similar to (4.32) can be derived for the error that occurs as a result of the discretization of the forward problem in electrical impedance tomography, optical fluorescence tomography, bioluminescence tomography, and microwave imaging.

2.4.3 Iterative Born Approximation

In this section, we explore the error in the inverse problem solution within an iterative linearization approach.

The error analysis presented in this paper covers the error which results from the discretization of the forward and inverse problems. If α is sufficiently low, then one iteration suffices to solve the inverse problem and the error analysis discussed above applies. When iterative linearization is considered to address the nonlinearity of the inverse problem, we can make use of the error analysis at each linearized step as follows: Let $\alpha^{\lambda}_{(t)}$ and $\tilde{\alpha}^{\lambda}_{n(t)}$ be the actual solution of the regularized inverse problem (5.19) and the solution of (2.26) at the t^{th} linearization step, respectively. At the end of the $(r-1)^{th}$ linearization step, the absorption coefficient estimate at $\mathbf{x}$ is given by $\hat{\mu}_a^{(r-1)}(\mathbf{x}) = \mu_a^{(0)}(\mathbf{x}) + \sum_{t=1}^{r-1} \tilde{\alpha}^{\lambda}_{n(t)}(\mathbf{x})$, where $\tilde{\alpha}^{\lambda}_{n(t)}$ has an error due to discretization with respect to the actual solution $\alpha^{\lambda}_{(t)}$, and $\mu_a^{(0)}$ is the initial guess for the background absorption coefficient. In the next linearization, an error on the new solution update $\hat{\mu}_a^{(r)}$ will be introduced due to:

1. projection (inverse problem discretization),

2. the error $(\tilde{\mathcal{K}} - \mathcal{K})^{(r-1)}$ in the operator $(\tilde{\mathcal{K}})^{(r-1)}$ and the error $(\tilde{\gamma} - \gamma)^{(r-1)}$ in $(\tilde{\gamma})^{(r-1)}$ resulting from the forward problem discretization, and

3. the error in the $(r-1)^{th}$ update $\hat{\mu}_a^{(r-1)}$, resulting from the discretization of the forward and inverse problems. Note that $\hat{\mu}_a^{(r-1)}$ appears as a coefficient

in the boundary value problems (5.1)-(5.2) and (5.3)-(5.4). An error in this coefficient implies perturbation in the solutions of (5.1)-(5.2) and (5.3)-(5.4). As a result, G_j and G_i^* will have error terms in addition to the discretization error.

As a result, the error in $\hat{\mu}_a^{(r)}$ at the r^{th} iteration is bounded by:

$$\|\mu_a - \hat{\mu}_a^{(r)}\| = \left\| \sum_{t=1}^{r} \alpha_{(t)}^{\lambda} - \tilde{\alpha}_{n(t)}^{\lambda} \right\| \leq \sum_{t=1}^{r} \|\alpha_{(t)}^{\lambda} - \tilde{\alpha}_{n(t)}^{\lambda}\|, \tag{2.63}$$

assuming that the initial guess $\mu_a^{(0)}$ for the background absorption is approximated accurately while solving the boundary value problems (5.1)-(5.2) and (5.3)-(5.4) at the first iteration, that is $\mu_a^{(0)}(\mathbf{x}) - \sum_{k=1}^{n} \mu_a^{(0)}(\mathbf{x}_k) L_k(\mathbf{x}) \to 0$, for all $\mathbf{x} \in \Omega$.

2.5 Conclusion

In this work, we presented an error analysis to show the relationship between the error in the reconstructed optical absorption images and the discretization of the forward and inverse problems. We summarized the implications of the error analysis in two theorems which provide an insight into the impact of forward and inverse problem discretizations on the accuracy of the reconstructed optical absorption images. These theorems show that the error in the reconstructed optical image due to the discretization of each problem is bounded by roughly the multiplication of the discretization error in the corresponding solution and the solution of the other problem. In particular, Theorem 2 shows that solving the diffusion equation and the associated adjoint problem accurately may not ensure small values for $\|\tilde{\mathcal{K}} - \mathcal{K}\|$ and $\|\gamma - \tilde{\gamma}\|$, which may lead to large errors in the reconstructed optical images, depending on the value of the regularization parameter. Similarly, relatively large discretization error in the solution of the forward problem may have relatively low impact on the accuracy of the reconstructed optical images, depending on the source-detector configuration, and orientation with respect to the optical heterogeneities. We have also shown that the error estimates can be extended to include the effect of noise on the overall error in the reconstructed images.

The error analysis presented in this work motivates the development of novel adaptive discretization schemes based on the error estimates in Theorems 1 and 2. In the sequel of this work, we propose two novel adaptive discretization algorithms for the forward and inverse problems [40], and justify the validity of Theorems 1 and 2.

The error analysis can be extended to show the effect of the discretization error on the accuracy of the simultaneous reconstruction of scattering and absorption coefficients, which will be the focus of our future work. Furthermore, the framework, which is used to analyze the effect of discretization error in the forward problem solution, can be employed to analyze the effect of perturbations in the kernel of the integral equation (4.18) on the accuracy of optical imaging. Finally, we note that the error analysis introduced in this paper is not limited to DOT, and can easily be adapted for similar inverse parameter estimation problems such as electrical impedance tomography, bioluminescence tomography, optical fluorescence tomography, microwave imaging etc.

CHAPTER 3

Adaptive mesh generation in diffuse optical absorption imaging

3.1 Introduction

Numerical approaches in solving the forward and inverse problems in Diffuse Optical Tomography (DOT) poses a tradeoff between computational efficiency and imaging accuracy. This tradeoff is a direct consequence of the discretization of the forward and inverse problems [39, 4] and the size of the resulting discrete forward and inverse problems. The imaging accuracy depends on the discretization error in the forward and inverse problem solutions. On the other hand, attempting to minimize the discretization error in the solutions of both problems separately implies a significant increase in the size of the discrete forward and inverse problems. Hence, it is important to understand the relationship between the discretization error and the resulting error in the solution of the inverse problem. Such a relationship can illuminate the mutual dependence of the forward and inverse problem solutions and identify the factors that control the extent to which the discretization error in the solutions of the forward and inverse problems affects the accuracy of the reconstructed optical images.

In part I of this two-part study, we present an error analysis which shows the effect of discretization of the forward and inverse problems on the accuracy of the reconstructed optical absorption images [39]. The analysis leads to two new error estimates that take into account the interdependence of the forward and inverse problems (see Section 3.2). In the second part of our work, based on the error analysis presented in part I, we develop new adaptive discretization schemes for the forward and inverse problems. The resulting locally refined meshes reduce the error in the reconstructed optical images while keeping the size of the discrete forward and inverse problems relatively small.

There has been extensive research on adaptive mesh generation for the numerical solution of partial differential equations (see [39] for a list of publications) and inverse parameter estimation problems to reduce the undesired effect of discretization error [19, 66]. In the area of DOT, in [5] it was numerically shown that approximation errors resulting from the discretization of the forward problem can lead to significant degradation in the quality of the reconstructed images. In that work, the error in the reconstructed images is minimized by using an enhanced imaging model that treats this additional approximation error within the Bayesian framework. Alternatively, several investigators have reported on adaptive discretization schemes for the forward and inverse problems to address the optical image degradation due to discretization. In [28] a "data driven zonation" scheme, which can be viewed as an adaptive discretization algorithm, was proposed for fluorescence imaging [28]. In [36], we presented a region-of-interest (ROI) imaging scheme for DOT, which employed a multi-level algorithm on a nonuniform grid. The non-uniform grid is designed so as to provide finer spatial resolution for the ROI which corresponds to the tumor region as indicated by *a priori* anatomical image. In [82] an *a priori* non-uniform mesh design which provides high resolution at the heterogeneities and near boundary regions was proposed. In that work, the mesh refinement is independent of the source-detector configuration and the location of the heterogeneities. In [35] a dual mesh strategy was proposed, in which, a relatively fine uniform mesh is considered for the forward problem discretization and a coarse uniform mesh is generated for the inverse problem discretization. In the same study, an adaptive refinement scheme was proposed for the inverse problem discretization, but no adaptive refinement was considered for the solution of the forward problem. Another dual mesh strategy which makes use of *a priori* ultrasound information was presented in [46]. In that work, the dual mesh is a coarse mesh for the background tissue and a relatively fine mesh for the heterogeneity, similar to the approach in [36]. In fluorescence imaging, a dual adaptive mesh strategy was used to discretize the inverse problem and the associated coupled diffusion equations, where the refinement criterion is based on *a posteriori* discretization error estimates [54]. Note that in all these studies [28, 36, 82, 35, 46, 54], the mesh refinement criteria considered for the inverse (forward) problem disregard the impact of the solution of the forward (inverse) problem. In other words, the discretization of each problem is considered independently of the solution of the other problem.

In this work, based on the two error bounds provided by the error analysis in part I [39], we introduce an adaptive discretization scheme for the forward and inverse problems, respectively. We remark that the mesh refinement criterion for each problem comprises the discretization error in the corresponding problem solution, scaled spatially by the solutions of both problems. Thus, the proposed adaptive mesh generation algorithms address the interdependence between the solutions of the forward and inverse problems and take into account the orientation of the source-detectors and the absorptive perturbations. This makes the adaptive discretization algorithms introduced in this paper different from the previous approaches [28, 36, 82, 35, 46, 54]. The simulation experiments validate the implications of our error analysis and show that the proposed mesh generation algorithms significantly improve the accuracy of the reconstructed optical images for a given number of unknowns in the discrete forward and inverse problems. We specifically show that using the discretization error estimates, which do not take into account the interdependence of forward and inverse problems as a criterion for discretization, may lead to severely degraded image reconstructions (see simulation study 3). We also discuss the computational complexity of the proposed adaptive mesh generation algorithms and compare it to the computational complexity of mesh generation algorithms based on the conventional discretization error estimates. We finally note that the proposed adaptive mesh generation algorithms can be adapted for similar inverse parameter estimation problems, such as electrical impedance tomography, optical fluorescence tomography, bioluminescence tomography, microwave imaging, etc.

The outline of this paper is as follows: In Section 2, we give a brief overview of the forward and inverse DOT problems and recall the two theorems presented in part I which summarize the impact of discretization on the accuracy of the reconstructed optical images. In Section 3, based on these two theorems, we introduce the adaptive mesh generation algorithms for the solution of the forward and inverse problems and discuss their computational complexity. In Section 4, we present our experimental results, which is followed by the Conclusion section. The Appendix includes the solution of a model problem used to initiate the adaptive mesh generation.

Table 3.1: Definition of function spaces and norms.

Notation	Explanation
$\overline{f}$	The complex conjugate of the function f
$C(\Omega)$	Space of continuous complex-valued functions on Ω
$C^k(\Omega)$	Space of complex-valued k-times continuously differentiable functions on Ω
$L^\infty(\Omega)$	$L^\infty(\Omega) = \{f \mid ess\sup_\Omega \lvert f(\mathbf{x})\rvert < \infty\ \}$
$L^p(\Omega)$	$L^p(\Omega) = \{f \mid (\int_\Omega \lvert f(\mathbf{x})\rvert^p d\mathbf{x})^{1/p} < \infty\ \}$, $p \in [1, \infty)$
$D_w^z f$	z^{th} weak derivative of f
$H^p(\Omega)$	$H^p(\Omega) = \{f \mid (\sum_{\lvert z\rvert \leq p} \lVert D_w^z f\rVert_0^2)^{1/2} < \infty\ \}$, $p \in [1, \infty)$
$\lVert f\rVert_0$	The $L^2(\Omega)$ norm of f
$\lVert f\rVert_p$	The $H^p(\Omega)$ norm of f
$\lVert f\rVert_\infty$	The $L^\infty(\Omega)$ norm of f
$\lVert f\rVert_{L^p(\Omega)}$	The $L^p(\Omega)$ norm of f
$\lVert f\rVert_{0,m}$	The L^2 norm of f over the m^{th} finite element Ω_m
$\lVert f\rVert_{p,m}$	The H^p norm of f over the m^{th} finite element Ω_m

3.2 Overview

In this section, we first briefly define the forward and inverse problems in DOT. Next, we state Theorems 1 and 2 presented in the first part of this work [39] to recall the effect of the discretization of the forward and inverse problems on the accuracy of optical absorption image reconstruction. We refer to Table 5.2 for the explanation of the notation associated with functions and their norms. Note that calligraphic letters are used to denote the operators, e.g. $\mathcal{A}_a$, $\mathcal{I}$, $\mathcal{K}$ etc.

3.2.1 Forward and inverse problems in DOT

We consider the following boundary value problem to model the near infrared light propagation in a bounded domain $\Omega \subset \mathbb{R}^3$ with Lipschitz boundary $\partial\Omega$ [20, 4]:

$$-\nabla \cdot D(\mathbf{x})\nabla g_j(\mathbf{x}) + (\mu_a(\mathbf{x}) + \frac{i\omega}{c})g_j(\mathbf{x}) = Q_j(\mathbf{x}) \quad \mathbf{x} \in \Omega, \tag{3.1}$$

$$g_j(\mathbf{x}) + 2aD(\mathbf{x})\frac{\partial g_j}{\partial n}(\mathbf{x}) = 0 \quad \mathbf{x} \in \partial\Omega, \tag{3.2}$$

where $g_j(\mathbf{x})$ is the photon density at $\mathbf{x}$, Q_j is the point source located at the source position $\mathbf{x}_s^j$, $D(\mathbf{x})$ is the diffusion coefficient and $\mu_a(\mathbf{x})$ is the absorption coefficient at $\mathbf{x}$, $i = \sqrt{-1}$, ω is the modulation frequency of the source, c is the speed of the light, $a = (1+R)/(1-R)$ where R is a parameter governing the internal reflection at the boundary $\partial\Omega$, and $\partial \cdot /\partial n$ denotes the directional derivative along the unit normal vector on the boundary. The boundary value problem (5.1)-(5.2) constitutes the forward problem in DOT together with the associated adjoint problem [4, 39]:

$$-\nabla \cdot D(\mathbf{x})\nabla g_i^*(\mathbf{x}) + (\mu_a(\mathbf{x}) - \frac{i\omega}{c})g_i^*(\mathbf{x}) = 0 \quad \mathbf{x} \in \Omega, \tag{3.3}$$

$$g_i^*(\mathbf{x}) + 2aD(\mathbf{x})\frac{\partial g_i^*}{\partial n}(\mathbf{x}) = Q_i^*(\mathbf{x}) \quad \mathbf{x} \in \partial\Omega, \tag{3.4}$$

where Q_i^* is the adjoint source located at the detector position $\mathbf{x}_d^i$. Note that we approximate the point source Q_j in (5.1) and the adjoint source Q_i^* in (5.4) by Gaussian functions with sufficiently low variance, whose centers are located at $\mathbf{x}_s^j$ and $\mathbf{x}_d^i$, respectively.

In this work, we focus on the estimation of the absorption coefficient and consider an iterative algorithm based on repetitive linearization of the inverse problem using first order Born approximation. Using a zeroth order Tikhonov regularization to address the illposedness, the inverse problem at each iteration reads:

$$\begin{aligned}\gamma(\mathbf{x}) = (\mathcal{A}_a^*\Gamma)(\mathbf{x}) &= [(\mathcal{A}_a^*\mathcal{A}_a + \lambda\mathcal{I})\alpha^\lambda](\mathbf{x}) \\ &:= \int_\Omega \kappa(\mathbf{x}, \acute{\mathbf{x}})\alpha^\lambda(\acute{\mathbf{x}})d\acute{\mathbf{x}} + \lambda\alpha^\lambda(\mathbf{x}) \qquad (3.5) \\ &:= (\mathcal{K}\alpha^\lambda)(\mathbf{x}), \qquad (3.6)\end{aligned}$$

where $\Gamma_{i,j}$ is the (i,j)th entry in the vector $\Gamma \in \mathbb{C}^{N_d \times N_s}$ which denotes the differential measurement at the i^{th} detector due to the j^{th} source as a result of the small perturbation α on the background absorption coefficient μ_a. In (3.5), $\kappa(\mathbf{x}, \acute{\mathbf{x}})$ is the kernel of the integral equation, given by [39]

$$\kappa(\mathbf{x}, \acute{\mathbf{x}}) = \sum_{i,j}^{N_d,N_s} H_{i,j}^*(\mathbf{x})H_{i,j}(\acute{\mathbf{x}}), \tag{3.7}$$

where $H_{i,j} = -\overline{g_i^*}g_j$ is the kernel of the integral operator $\mathcal{A}_a : L^\infty(\Omega) \to \mathbb{C}^{N_d \times N_s}$ and

$H^*_{i,j} = -g^*_i \overline{g_j}$ is the kernel of the adjoint operator $\mathcal{A}^*_a : \mathbb{C}^{N_d \times N_s} \to L^1(\Omega)$ given by:

$$(\mathcal{A}^*_a \beta)(\mathbf{x}) = \sum_{i,j}^{N_d,N_s} H^*_{i,j}(\mathbf{x})\beta_{i,j} = \sum_{i,j}^{N_d,N_s} -g^*_i(\mathbf{x})\overline{g_j(\mathbf{x})}\beta_{i,j}, \tag{3.8}$$

for all $\beta \in \mathbb{C}^{N_d \times N_s}$. Note that g_j and g^*_i in (4.18) and (5.15) are the solutions of the variational formulations of (5.1)-(5.2) and (5.3)-(5.4), respectively [39]. For the rest of the paper, we will denote $L^\infty(\Omega)$ and $L^1(\Omega)$ by X and Y, respectively.

Below we summarize the two theorems of part I [39] and provide the error estimates which will be used in the design of adaptive meshes for the discretization of the forward and inverse DOT problems. In this respect, we first consider the impact of inverse problem discretization when the associated kernel $\kappa(\mathbf{x}, \acute{\mathbf{x}})$ in (3.5) is exact. Next, we give the error estimate for the case in which the kernel is replaced by its finite dimensional approximation (i.e. degenerate kernel) and analyze the effect of the forward problem discretization on the accuracy of the reconstructed image without projecting (3.6).

3.2.2 Effect of inverse problem discretization

Consider the discretization of the inverse problem (3.6) by projecting it onto a finite dimensional subspace $Y_n \subset Y$ spanned by the first order Lagrange basis functions $\{L_1, \ldots, L_n\}$, using the collocation method [39]:

$$\mathcal{P}_n \mathcal{K} \alpha^\lambda_n = \mathcal{P}_n \gamma, \tag{3.9}$$

where $\mathcal{P}_n : Y \to Y_n$ is the projection operator associated with the collocation method with first order Lagrange basis functions [39] such that $\alpha^\lambda_n(\mathbf{x})$, $\mathbf{x} \in \Omega$, is approximated on a set $\{\Omega_m\}$ of finite elements for $m = 1, \cdots, N_\Delta$, $\bigcup_m^{N_\Delta} \Omega_m = \Omega$ as follows:

$$\alpha^\lambda_n(\mathbf{x}) = \sum_{k=1}^{n} a_k L_k(\mathbf{x}). \tag{3.10}$$

Let ψ be the interpolant of α^λ [20] and assume that $\alpha^\lambda \in H^1(\Omega)$. Then, the interpolation error $e_\alpha = \alpha^\lambda - \psi$ on each finite element is bounded by

$$\|e_\alpha\|_{0,m} \leq C\|\alpha^\lambda\|_{1,m}h_m, \tag{3.11}$$

where C is a positive constant and h_m is the diameter of the smallest ball that contains the m^{th} element.

Theorem 1 describes the effect of inverse problem discretization on the accuracy of the reconstructed optical absorption image.

Theorem 1: The error between the solution α^λ of (3.6) and the solution α_n^λ of (3.9) is bounded by:

$$\|\alpha^\lambda - \alpha_n^\lambda\|_{L^1(\Omega)} \leq C\sqrt{V_\Omega}\|\mathcal{I} - \mathcal{T}_n\|_{Y\to X_n} \sum_{m=1}^{N_\Delta} \|\alpha^\lambda\|_{1,m}h_m$$

$$+\frac{C}{\lambda}\|\mathcal{T}_n\|_{Y\to X_n} \max_{i,j} \|g_i^* g_j\|_{L^1(\Omega)} \sum_{m=1}^{N_\Delta} \sum_{i,j}^{N_d,N_s} \|g_i^* g_j\|_{0,m}\|\alpha^\lambda\|_{1,m}h_m, \tag{3.12}$$

where C is a positive constant, V_Ω is the volume of Ω, $\mathcal{T}_n : Y \to X_n$ is a uniformly bounded operator given by $\mathcal{T}_n = (\mathcal{I} + \frac{1}{\lambda}\mathcal{P}_n\mathcal{A}_a^*\mathcal{A}_a)^{-1}\mathcal{P}_n$, and g_j, g_i^* are the solutions of the variational formulations of the boundary value problems (5.1)-(5.2) and (5.3)-(5.4), respectively [39].

Proof: See [39].

3.2.3 Effect of forward problem discretization

Assume that $D, \mu_a \in C^1(\Omega)$. Noting that $Q_j, Q_i^* \in H^1(\Omega)$, the solutions g_j, g_i^* of the variational formulations of the boundary value problems (5.1)-(5.2) and (5.3)-(5.4) satisfy [39]

$$g_j, g_i^* \in C(\Omega). \tag{3.13}$$

Let $\{\Omega_m^j\}$ denote the set of linear elements used to discretize the variational formulation of the boundary value problem (5.1)-(5.2) for $m = 1, \ldots, N_\Delta^j$; such that $\bigcup_m^{N_\Delta^j} \Omega_m^j = \Omega$ and h_m^j be the diameter of the smallest ball that contains the m^{th}

element in the finite dimensional solution G_j, for all $j = 1, \ldots N_s$ [39]. Similarly, let $\{\Omega_n^i\}$ denote the set of linear elements used to discretize the variational formulation of the boundary value problem (5.3)-(5.4) for $n = 1, \ldots, N_\Delta^{*i}$; such that $\bigcup_n^{N_\Delta^{*i}} \Omega_n^i = \Omega$ and h_n^i be the diameter of the smallest ball that contains the n^{th} element in the finite dimensional solution G_i^*, for all $i = 1, \ldots N_d$ [39]. Then, a bound for the discretization error in the finite element solutions G_j and G_i^* with respect to the solutions g_j and g_i^* of the variational formulations of the boundary value problems (5.1)-(5.2) and (5.3)-(5.4) on each finite element can be given by [20]

$$\|g_j - G_j\|_{0,m^j} \leq C\|g_j\|_{1,m^j} h_m^j, \tag{3.14}$$

$$\|g_i^* - G_i^*\|_{0,n^i} \leq C\|g_i^*\|_{1,n^i} h_n^i, \tag{3.15}$$

where C is a positive constant, and $\|\cdot\|_{0,a^b}$ and $\|\cdot\|_{1,a^b}$ denote respectively the L^2 and H^1 norms on the finite element Ω_a^b.

Consider the inverse problem

$$\tilde{\mathcal{K}}\tilde{\alpha}^\lambda = \tilde{\gamma}, \tag{3.16}$$

where $\tilde{\mathcal{K}}$ and $\tilde{\gamma}$ are the finite dimensional approximations to $\mathcal{K}$ and γ, obtained by substituting g_j and g_i^* in $H_{i,j}$ and $H_{i,j}^*$ by G_j and G_i^*, respectively.

Theorem 2 shows the effect of forward problem discretization on the accuracy of the reconstructed optical absorption image.

Theorem 2: A bound for the error between the solution α^λ of (3.6) and the solution $\tilde{\alpha}^\lambda$ of (3.16) due to approximations $\tilde{\mathcal{K}}$ and $\tilde{\gamma}$ is given by:

$$\begin{aligned} \|\alpha^\lambda - \tilde{\alpha}^\lambda\|_{L^1(\Omega)} \leq \frac{C}{\lambda} \max_{i,j} \|g_i^* g_j\|_{L^1(\Omega)} \\ \times \left(\sum_{i=1}^{N_d} \sum_{n,j}^{N_\Delta^{*i}, N_s} \Big(2\|g_j \alpha^\lambda\|_{0,n^i} + \|\alpha\|_\infty \|g_j\|_{0,n^i} \Big) \|g_i^*\|_{1,n^i} h_n^i \right. \\ \left. + \sum_{j=1}^{N_s} \sum_{m,i}^{N_\Delta^j, N_d} \Big(2\|g_i^* \alpha^\lambda\|_{0,m^j} + \|\alpha\|_\infty \|g_i^*\|_{0,m^j} \Big) \|g_j\|_{1,m^j} h_m^j \right), \end{aligned} \tag{3.17}$$

where C is a positive constant.

Proof: See [39].

We refer to part I [39] for further details of the discussion regarding the definition and discretization of the forward and inverse problems. In the following, we discuss the adaptive mesh generation for the forward and inverse problems.

3.3 Adaptive Mesh Generation

In this section, we discuss the adaptive mesh design for the discretization of the forward and inverse problems based on the Theorems 1 and 2. For each problem, we present an adaptive mesh generation algorithm, which is followed by the corresponding computational cost analysis.

3.3.1 Adaptive Mesh Generation for the Forward Problem

Let the mesh parameter h_m^j for G_j for $j = 1, \cdots, N_s$ and the mesh parameter h_m^i for G_i^* for $i = 1, \cdots, N_d$ be chosen so that:

$$h_m^j \leq \frac{\epsilon_f}{\sum_i^{N_d}(2\|g_i^* \alpha^\lambda\|_{0,m^j} + \|\alpha\|_\infty \|g_i^*\|_{0,m^j})\|g_j\|_{1,m^j}} := B_j^m, \tag{3.18}$$

$$h_n^i \leq \frac{\epsilon_f}{\sum_j^{N_s}(2\|g_j \alpha^\lambda\|_{0,n^i} + \|\alpha\|_\infty \|g_j\|_{0,n^i})\|g_i^*\|_{1,n^i}} := B_i^{*n}, \tag{3.19}$$

where the tolerance ϵ_f will be defined later. Then, by Theorem 2, the error in the reconstructed image due to the forward problem discretization is bounded by:

$$\frac{C}{\lambda} \max_{i,j} \|g_i^* g_j\|_{L^1(\Omega)} \left(\sum_j^{N_s} N_\Delta^j + \sum_i^{N_d} N_\Delta^{*i} \right) \epsilon_f = \tilde{\epsilon}^f, \tag{3.20}$$

where C is a positive constant and $\tilde{\epsilon}^f$ is the total allowable error in the reconstructed optical image due to the forward problem discretization. Equation (3.20) implies the following value for ϵ_f:

$$\epsilon_f = \frac{\lambda \tilde{\epsilon}^f / C}{\max_{i,j} \|g_i^* g_j\|_{L^1(\Omega)}} \frac{1}{\left(\sum_j^{N_s} N_\Delta^j + \sum_i^{N_d} N_\Delta^{*i}\right)}. \tag{3.21}$$

Algorithm 1 outlines the adaptive mesh generation algorithm for the forward problem in the form of a pseudocode. The algorithm is performed for each source and detector before the linearization of the inverse problem and it yields a family of adaptively refined meshes with conforming elements. We use Rivara's algorithm [76] for refinement.

Algorithm 1 The pseudocode for the mesh generation algorithm for the forward problem, prior to the linearization of the inverse problem.

```
◇ Generate an initial uniform mesh $(\Delta,N_\Delta)$, $\Delta = \bigcup_{m=1}^{N_\Delta}\{\Delta_m\}$
◇ Set $\epsilon^f$
◇ Initialize the set of marked elements: $M_e \leftarrow \{\}$
◇ flag = True
  while flag = True
    for each element $\Delta_m \in \Delta$ with mesh parameter $h_m^j$ $(h_m^i)$
      if first linearization
        . Use analytical solutions for $g_j$ and $g_i^*$ and a priori anatomical
        information about $\alpha$ to compute the bound $B_j^m$ in (3.18) ($B_i^{*n}$ in (3.19))
      else
        . Use current solution updates $G_j$ and $G_i^*$ and $\tilde{\alpha}_n^\lambda$
        to compute $B_j^m$ in (3.18) ($B_i^{*n}$ in (3.19))
      end
      if $h_m^j > B_j^m$ $(h_m^i > B_i^{*n})$
        . $M_e \leftarrow M_e \bigcup \{\Delta_m\}$
      end
    end

    if $M_e \neq \{\}$
      . Refine the marked elements and update the mesh $\Delta$
      . $M_e \leftarrow \{\}$
    else
      . flag = False
    end
  end
◇ Solve for $G_j$ $(G_i^*)$.
```

Remarks 1:

1. In practice, B_j^m and B_i^{*n} in (3.18)-(3.19) can not be computed since α, α^λ, g_j, and g_i^* are unknown. However, B_j^m and B_i^{*n} can be estimated by using approximations for the functions involved in these bounds, based on either

a priori information or on the recent forward and inverse problem solution updates. Then, the elements whose mesh parameter h^j_m (h^i_n) exceeds B^m_j (B^{*n}_i) can be determined and refined.

2. After the first sweep of refinement, one can compute the bound B^m_j and B^{*n}_i only for the new elements. We note that for the initial mesh design, we use a model problem to compute the terms in the error bound relevant to the forward problem solution (see B.1). If there is no *a priori* information, α^λ can be assumed to be spatially constant at the first linearization step. After the first linearization, the norms in B^m_j and B^{*n}_i relevant to g_j and g^*_i are not expected to change significantly. In this context, the terms $\|g^*_i\alpha^\lambda\|_{0,m}$, $\|g_j\alpha^\lambda\|_{0,n}$ in (3.18) and (3.19) can be bounded by $\|g^*_i\|_{0,m^j}\|\alpha^\lambda\|_{\infty,m^j}$ and $\|g_j\|_{0,n^i}\|\alpha^\lambda\|_{\infty,n^i}$, respectively. Therefore, one can store the norms $\|g_j\|_{0,m^j}$ and $\|g^*_i\|_{0,n^i}$ at the end of the first mesh generation, and update B^m_j and B^{*n}_i in the following mesh generations by using these stored values and the updated α^λ values.

3. In case ϵ_f can not be chosen in prior, we consider a posterior approach, set $\epsilon_f = 1$, and compute h^j_m/B^m_j (h^i_n/B^{*n}_i) on each element, which is used as the indicator for refinement. Then, the elements with indicator value exceeding the average h^j_m/B^m_j (h^i_n/B^{*n}_i) quantity are marked for refinement. We note that in this case, the algorithm has to be stopped when the number of nodes in the mesh exceeds the allowable number of nodes.

3.3.2 Computational cost of the adaptive mesh generation algorithm for the forward problem

Consider the algorithm described in Remark 1(iii) for $\Omega \subset \mathbb{R}^2$. Using triangular finite elements with first order Lagrange basis functions and an analytical (exact) integration on each finite element, the number of multiplications required to compute the L^2 or H^1 norm of a finite dimensional function on each triangular element Ω^j_m (Ω^i_n) is 12. On the other hand, computing the norm $\|g_j\alpha^\lambda\|_{0,n^i}$ ($\|g^*_i\alpha^\lambda\|_{0,m^j}$) takes 10 times the number of multiplications to compute $\|g_j\|_{1,m^j}$ ($\|g^*_i\|_{1,n^i}$). As a result, the total number of multiplications required to compute the error estimates on all finite elements for the j^{th} source is given by $(132N_d + 16)N^j_\Delta$. Similarly, the total number of multiplications required to compute the error estimates on all finite elements for the i^{th} detector is equal to $(132N_s + 16)N_\Delta^{*i}$.

In order to reduce the computational cost of the proposed adaptive mesh generation algorithm, we can approximate the bounds B_j^m in (3.18) and B_i^{*n} in (3.19) as follows:

$$B_j^m \approx \frac{\epsilon_f}{(2\|\sum_i^{N_d} g_i^* \alpha^\lambda\|_{0,m^j} + \|\alpha\|_\infty \|\sum_i^{N_d} g_i^*\|_{0,m^j})\|g_j\|_{1,m^j}} \tag{3.22}$$

$$B_i^{*n} \approx \frac{\epsilon_f}{(2\|\sum_j^{N_s} g_j \alpha^\lambda\|_{0,n^i} + \|\alpha\|_\infty \|\sum_j^{N_s} g_j\|_{0,n^i})\|g_i^*\|_{1,n^i}}. \tag{3.23}$$

Then, the number of multiplications required to compute the error estimates on all finite elements becomes $148N_\Delta^j$ $(148N_\Delta^{*i})$, which implies a significant reduction as compared to $(132N_d + 16)N_\Delta^j$ $((132N_s + 16)N_\Delta^{i*})$.

If one uses the discretization error estimates (3.14)-(3.15) to generate adaptive meshes for the discretization of (5.1)-(5.2) and (5.3)-(5.4), the number of multiplications is equal to $13N_\Delta^j$ and $13N_\Delta^{*i}$, respectively. Then, the resulting adaptive meshes will lead to finite element solutions G_j and G_i^* with reduced discretization error. However, reduction in the discretization error in G_j and G_i^* may not ensure the accuracy of the reconstructed absorption image (see simulation experiment 3).

3.3.3 Adaptive mesh generation for the inverse problem:

Let the mesh parameter h_m for the solution of the inverse problem be defined as follows:

$$\begin{aligned} h_m \leq \epsilon_{inv} / \Bigg(& \sqrt{V_\Omega}\|\mathcal{I} - \mathcal{T}_n\|_{Y \to X_n}\|\alpha^\lambda\|_{1,m} + \frac{1}{\lambda}\|\mathcal{T}_n\|_{Y \to X_n} \\ & \times \max_{i,j} \|g_i^* g_j\|_{L^1(\Omega)} \sum_{i,j}^{N_d,N_s} \|g_i^* g_j\|_{0,m}\|\alpha^\lambda\|_{1,m} \Bigg) := B_{inv}^m. \end{aligned} \tag{3.24}$$

Then, by Theorem 1, the error in the reconstructed image due to inverse problem discretization is bounded by

$$C N_\Delta \epsilon_{inv} = \tilde{\epsilon}^{inv}, \tag{3.25}$$

where C is a positive constant and $\tilde{\epsilon}^{inv}$ is the total allowable error in the reconstructed optical image due to inverse problem discretization.

We present the pseudocode for our adaptive mesh generation algorithm used at each linearization of the inverse problem in Algorithm 2. Similar to the forward problem discretization, we use Rivara's algorithm [76] for the refinement of the elements.

Algorithm 2 The pseudocode for the mesh generation algorithm at every linearization step of the inverse problem.

```
◇ Generate an initial uniform mesh $(\Delta, N_\Delta)$, $\Delta = \bigcup_{m=1}^{N_\Delta}\{\Delta_m\}$
◇ Set $\epsilon_{inv}$
◇ Initialize the set of marked elements: $M_e \leftarrow \{\}$
◇ flag = True
  while flag = True
    for each element $\Delta_m \in \Delta$ with mesh parameter $h_m$
      if first linearization
        . Use current solution updates $G_j$ and $G_i^*$ and a priori information
          about $\alpha$ to compute $B_{inv}^m$ in (3.24)
      else
        . Use current solution updates $G_j$ and $G_i^*$ and $\tilde{\alpha}_n^\lambda$
          to compute $B_{inv}^m$ in (3.24)
      end
      if $h_m > B_{inv}^m$
        . $M_e \leftarrow M_e \bigcup \{\Delta_m\}$
      end
    end

    if $M_e \neq \{\}$
      . Refine the marked elements and update the mesh $\Delta$
      . $M_e \leftarrow \{\}$
    else
      . flag = False
    end
  end
◇ Solve for $\tilde{\alpha}_n^\lambda$.
```

Remarks 2:

1. In practice, B_{inv}^m in (3.24) can not be computed since α^λ, g_j, g_i^*, and $\mathcal{T}_n$ are unknown. Similar to the approach described in Section 3.3.1, we can

compute an estimate for B_{inv}^m by using the uniform boundedness of the operator $\mathcal{T}_n$ [39] and by using approximate values for the functions involved in B_{inv}^m. In this context, we use either *a priori* information or the recent forward and inverse problem solution updates to calculate (3.24) on each element. Then, the elements with the mesh parameter $h_m > B_{inv}^m$ are determined and refined.

2. In order to save computations, after the first sweep of refinement, one can compute the bound B_{inv}^m only for the new elements. Furthermore, similar to the approach described in Section 3.3.1, the term $\|g_i^* g_j\|_{0,m}$ in (3.24) can be stored after the first mesh generation and can be used in the following mesh generations. In this context, the bound B_{inv}^m can be updated by using only the updated $\|\alpha^\lambda\|_{1,m}$ value.

3. Note that, in practice, one of the two terms in the denominator of B_{inv}^m will be dominant depending on the value of λ. Thus, we consider only the dominant term for the computation of B_{inv}^m. In case ϵ_f can not be chosen in prior, we consider a posterior approach, set $\epsilon_{inv} = 1$ and compute $\|\alpha^\lambda\|_{1,m} h_m$ or $\sum_{i,j}^{N_d,N_s} \|g_i^* g_j\|_{0,m} \|\alpha^\lambda\|_{1,m} h_m$ on each element, which are used as refinement indicators. Then, the elements with indicator value which exceeds the average indicator value are refined. In this case, the algorithm has to be stopped when the number of nodes in the mesh exceeds the allowable number of nodes.

3.3.4 Computational cost of the adaptive mesh generation algorithm for the inverse problem

Consider the algorithm stated in Remark 2(iii) for $\Omega \subset \mathbb{R}^2$ and assume that the second term in the denominator of B_{inv}^m (3.24) is dominant. Using triangular finite elements with first order Lagrange basis functions and an analytical (exact) integration on each finite element, the total number of multiplications required to compute the error estimates on all finite elements is given by $(120 N_d N_s + 14) N_\Delta$.

In order to reduce the number of multiplications, we can consider an approximation for B^m_{inv} as follows:

$$B^m_{inv} \approx \frac{1}{\left\| \sum_{i,j}^{N_d,N_s} g_i^* g_j \right\|_{0,m} \|\alpha^\lambda\|_{1,m}}. \tag{3.26}$$

Then, the number of multiplications reduces to $134N_\Delta$.

If one uses the interpolation error estimate (3.11) to generate adaptive meshes, the number of multiplications to compute the error estimates on all finite elements will be $13N_\Delta$. However, such adaptive meshes may not help reduce the error in the reconstructed optical images, resulting from discretization (see simulation experiment 3).

3.4 Numerical Experiments

We conduct a series of numerical experiments to demonstrate the implications of Theorems 1 and 2, and to present the effectiveness of the proposed adaptive mesh generation algorithms. We perform our experiments in 2D for ease of comparison.

In the first simulation, we consider a series of image reconstructions to show the effectiveness of the proposed adaptive mesh generation algorithms. In this context, we compare the images reconstructed by using uniform meshes for the forward and inverse problems to the images reconstructed by using adaptive meshes which are designed based on Theorems 1 and 2.

In the second simulation, we show the effect of the heterogeneity size on the accuracy of the reconstructed absorption images. Next, we demonstrate how this error can be addressed by the proposed adaptive discretization schemes.

In the final simulation study, we demonstrate the implication of Theorem 2 and show that meshes generated for the forward problem by using discretization error estimates which disregard the interaction between the solutions g_j and g_i^* and α^λ can lead to unstable image reconstructions. We note that the proposed adaptive mesh generation algorithm for the forward problem addresses this problem.

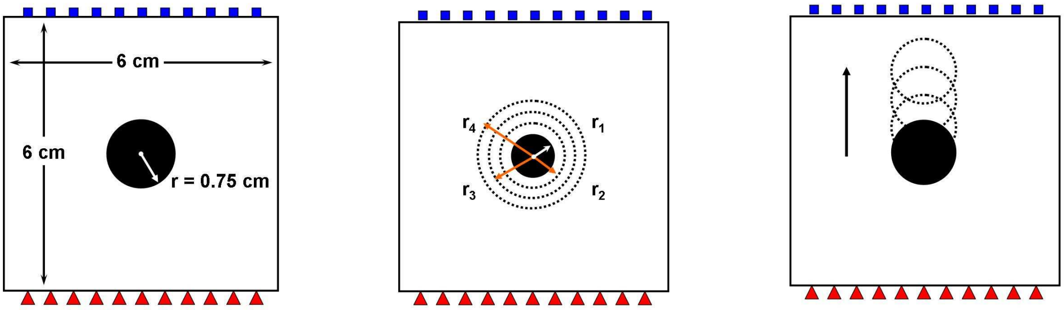
Note that in all experiments, we use triangular finite elements with first order Lagrange basis functions. We apply Gaussian elimination method to solve the discrete forward problem resulting from the variational formulation [20] of the boundary value problems (5.1)-(5.2) and (5.3)-(5.4) [39]. For the inverse problem, we consider the discrete problem obtained by projecting (3.16) by collocation method:

$$\mathcal{P}_n\tilde{\mathcal{K}}\tilde{\alpha}_n^\lambda = \mathcal{P}_n\tilde{\gamma}, \tag{3.27}$$

where the regularization parameter is chosen as small as possible, yet large enough to enable robust image reconstructions. In this respect, an appropriate value for the regularization parameter is chosen based on experience.

3.4.1 Simulation Study 1

In this simulation study, we consider the geometry shown in Figure 4.1(a). We simulate the optical data by solving the diffusion equation at $\omega = 0$ on a fine uniform grid with 61 nodes along x and y directions, where the refractive index mismatch parameter $a = 3$. 11 sources and 11 detectors are evenly spaced on the bottom and top edges of the square, respectively. The diffusion coefficient $D(\mathbf{x}) = 0.0410$ for $\mathbf{x} \in \Omega \cup \partial\Omega$. The circular heterogeneity with absorption coefficient $\mu_a = 0.2$ cm^{-1} is embedded in an optically homogeneous background with $\mu_a = 0.04$ cm^{-1}.



(a) The optical domain and source-detector configuration for simulation study 1.

(b) The optical domain and source-detector configuration for simulation study 2. $r_1 = 0.50$ cm, $r_2 = 0.75$ cm, $r_3 = 1.0$ cm, and $r_4 = 1.25$ cm.

(c) The optical domain and source-detector configuration for simulation study 3. The radius of the circles is 0.75 cm.

Figure 3.1: The setups used for the simulation studies 1, 2, and 3. The squares and triangles denote the detectors and sources, respectively.

In order to obtain a series of absorption imaging problems using the same setup, we consider 5 values for the background absorption value. Then, for each imaging problem, we consider three mesh scenarios: Uniform mesh for both forward and inverse problems; adaptive mesh for the forward problem and uniform mesh for the inverse problem; and adaptive meshes for both forward and inverse problems. We refer to Table 3.2 for a brief outline of the first simulation study.

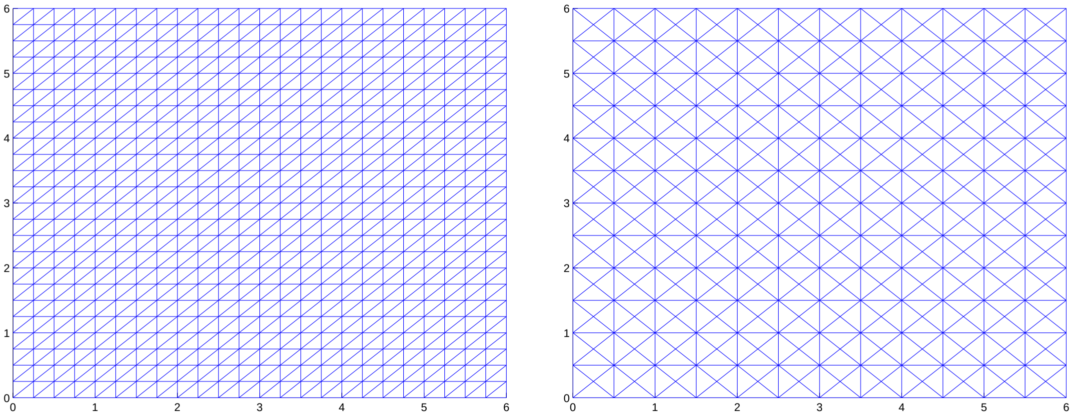
Table 3.2: The mesh scenarios and the background μ_a values in simulation study 1.

Mesh (Forward)	**Mesh (Inverse)**	Background μ_a (cm^{-1})
Uniform	Uniform	0.032, 0.036, 0.040, 0.044, 0.050
Adaptive	Uniform	0.032, 0.036, 0.040, 0.044, 0.050
Adaptive	Adaptive	0.032, 0.036, 0.040, 0.044, 0.050

The uniform mesh used for the forward problem discretization has 625 nodes and is shown in Figure 3.2(a). The uniform mesh for the inverse problem has 313 nodes and is shown in Figure 3.2(b). We use the algorithms described in Section 3.3.1 and Remark 1(iii), and Section 3.3.3 Remark 2(iii) to generate the adaptive meshes for the forward and inverse problems, respectively. The number of nodes in each of the adaptive meshes used for the forward problem does not exceed 750. An example for the adaptive mesh generated for a source located at $(1.0, 0)$ is shown in Figure 3.2(c). The adaptive mesh for the inverse problem generated for the case where the background $\mu_a = 0.050$ cm^{-1} has 418 nodes and is shown in Figure 3.2(d).

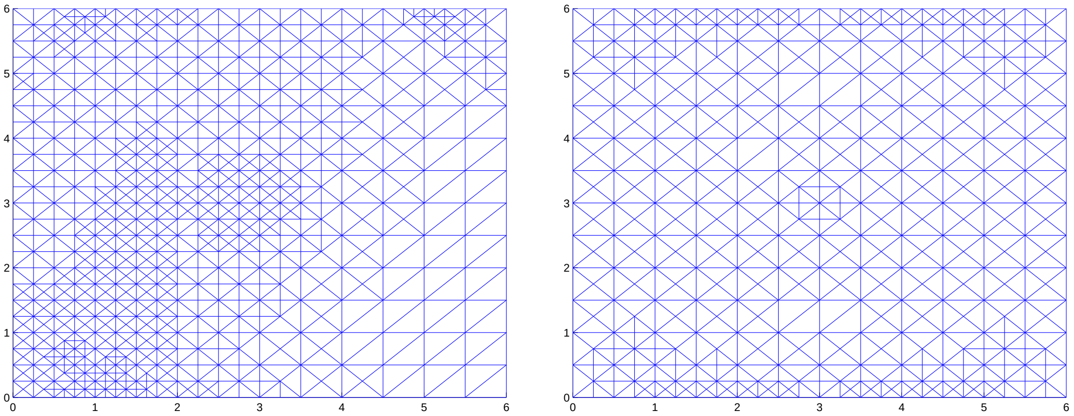
For the inverse problem, we set the regularization parameter λ to 10^{-7} in all experiments to eliminate the dependence of the error estimates (3.12)-(4.34) on the regularization parameter. We consider the image reconstructed by using fine uniform meshes (61×61 nodes for the forward problem and 61×61 nodes for the inverse problem) as the reference image, which is assumed to possess no error due to discretization. We compute the error $\|\alpha^\lambda - \tilde{\alpha}_n^\lambda\|_{L^1(\Omega)}$ for each image reconstruction and tabulate the results in Table 3.3. We see that the error in the images reconstructed by using uniform meshes for both forward and inverse problems is significantly reduced by the use of adaptively refined meshes. A similar behavior is observed for all choices of background absorption value.

We present image reconstructions in Figures 3.3 and 3.4 for the two extreme



(a) The uniform mesh with 625 nodes.

(b) The uniform mesh with 313 nodes.



(c) The adaptive mesh generated for the forward problem for the source located at (1.0,0): Background $\mu_a = 0.050$ cm^{-1}.

(d) The adaptive mesh generated for the inverse problem solution, with 418 nodes. Background $\mu_a = 0.050$ cm^{-1}.

Figure 3.2: Examples of meshes used in the first simulation study.

cases, where the background absorption value is equal to 0.032 and 0.050 cm^{-1}, respectively. Figures 3.3(a) and 3.4(a) display the reference images used to compute the error values given in Table 3.3. Figures 3.3(c) and 3.3(d) show that the optical heterogeneity is resolved better by using adaptive meshes as compared to the reconstructed image obtained by using uniform meshes, which is shown in Figure 3.3(b). These results are consistent with the error values given in Table 3.3. A similar trend is seen in Figures 3.4(c) and 3.4(d). Note that the number of nodes in the adaptive meshes is almost equal to the number of nodes that the uniform meshes have. In Figure 3.5, we show the cross-sectional views from the reconstructed images. We see that the use of coarse uniform meshes fails to resolve the circular heterogeneity

Table 3.3: The error $\|\alpha^\lambda - \tilde{\alpha}_n^\lambda\|_{L^1(\Omega)}$ for each experiment described in the simulation study 1 and Table 3.2. The first column shows the type of the meshes used in the forward and inverse problems, respectively. The unit of background μ_a is cm^{-1}.

	Background μ_a:	0.032	0.036	0.040	0.044	0.050
Uniform-Uniform	$\|\alpha^\lambda - \tilde{\alpha}_n^\lambda\|_{L^1(\Omega)}$:	0.2325	0.2559	0.2773	0.2932	0.3013
Adaptive-Uniform	$\|\alpha^\lambda - \tilde{\alpha}_n^\lambda\|_{L^1(\Omega)}$:	0.1238	0.1139	0.1166	0.1209	0.1278
Adaptive-Adaptive	$\|\alpha^\lambda - \tilde{\alpha}_n^\lambda\|_{L^1(\Omega)}$:	0.1043	0.0997	0.0998	0.1003	0.1009

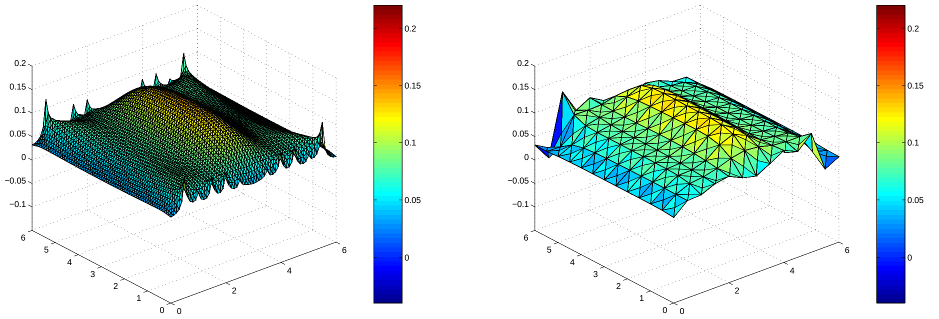
especially for the case in which the background $\mu_a = 0.032$ cm^{-1}.

3.4.2 Simulation Study 2

In this study, we consider the geometry shown in Figure 3.1(b). To simulate the optical data, we use the same source-detector configuration considered in the first simulation study. We simulate the optical data by solving the diffusion equation at $\omega = 0$ on a fine uniform grid with 61 nodes along x and y directions, where the refractive index mismatch parameter $a = 3$. The diffusion coefficient D is assumed to be constant and $D(\mathbf{x}) = 0.0410$ cm, for all $\mathbf{x} \in \Omega \cup \partial\Omega$.

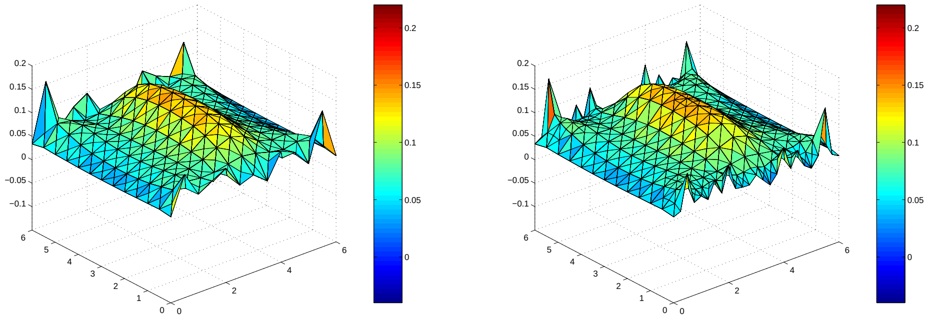
We consider 4 different radii for the circular heterogeneity with $\mu_a = 0.20$ cm^{-1} embedded in a background with $\mu_a = 0.040$ cm^{-1} as shown in Figure 3.1(b). For each case, we compute the error for different mesh scenarios, similar to the first simulation study: Uniform mesh for both forward and inverse problems; adaptive mesh for the forward problem and uniform mesh for the inverse problem; and adaptive meshes for both forward and inverse problems. The adaptive meshes for this simulation study were generated based on Theorems 1 and 2, and the mesh generation algorithms described in the first simulation study and Section 3.3. The uniform meshes used for the forward and inverse problems are identical to those used in the first simulation study. We note that the number of nodes in the adaptive meshes generated for the forward and inverse problems is close to the number of nodes in the corresponding uniform meshes.

In Table 3.4, we tabulate the error norm $\|\alpha^\lambda - \tilde{\alpha}_n^\lambda\|$ obtained for each heterogeneity size with different mesh choices. Table 3.4 shows that the error increases with increasing heterogeneity size. We see that the reduction in the error as a result



(a) The optical absorption image used as the reference for error computations.

(b) The reconstructed absorption image using the uniform mesh in Figure 3.2(a) for the forward, and the uniform mesh in Figure 3.2(b) for the inverse problem.



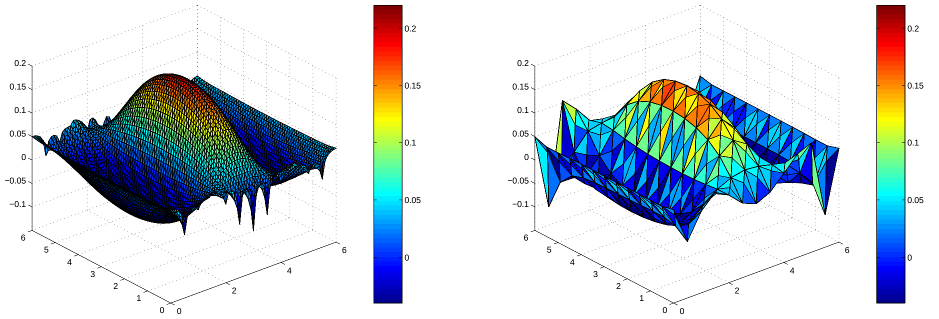
(c) The reconstructed absorption image using an adaptive mesh for the forward, and the uniform mesh in Figure 3.2(b) for the inverse problem.

(d) The reconstructed absorption image using an adaptive mesh for the forward, and the adaptive mesh in Figure 3.2(d) for the inverse problem.

Figure 3.3: The reconstruction results of simulation study 1, with the background $\mu_a = 0.032$ cm^{-1}.

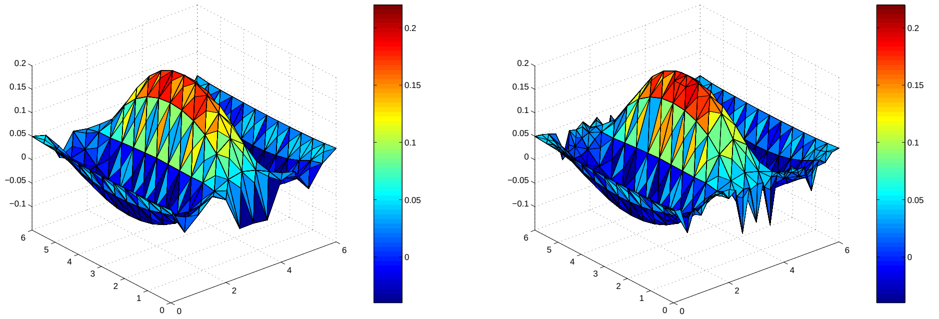
of using adaptive meshes is more significant for smaller sized heterogeneities. Further reduction in the error norm $\|\alpha^\lambda - \tilde{\alpha}_n^\lambda\|$ is possible by increasing the number of nodes in the meshes.

For brevity, we only show the reconstruction results for the extreme cases: $r = 0.5$ cm and $r = 1.25$ cm. We note that the regularization parameter $\lambda = 5 \times 10^{-9}$ in all reconstructions. Figures 3.6(a) and 3.6(b) show the images used as the reference images α^λ in the calculation of the error norms $\|\alpha^\lambda - \tilde{\alpha}_n^\lambda\|$ listed in Table 3.4. Figures 3.6(e)-3.6(f) show that the adaptive meshes reduce the artifacts as compared to the images reconstructed by using uniform meshes, which are shown in Figures 3.6(c)-3.6(d).



(a) The optical absorption image used as the reference for error computations.

(b) The reconstructed absorption image using the uniform mesh in Figure 3.2(a) for the forward, and the uniform mesh in Figure 3.2(b) for the inverse problem.



(c) The reconstructed absorption image using an adaptive mesh for the forward, and the uniform mesh in Figure 3.2(b) for the inverse problem.

(d) The reconstructed absorption image using an adaptive mesh for the forward, and the adaptive mesh in Figure 3.2(d) for the inverse problem.

Figure 3.4: The results of simulation study 1, with the background $\mu_a = 0.050$ cm^{-1}.

3.4.3 Simulation Study 3

In this simulation study, we consider the geometry shown in Figure 3.1(c). The center of the circular heterogeneity is moved vertically towards the detector side to see the effect on the imaging accuracy. Next, we show how the error in the reconstructed images due to discretization can be addressed by using appropriate meshes for the solutions of the forward and inverse problems. In this context, we compare the results obtained by using 1) uniform meshes, 2) the adaptive meshes generated using conventional *a priori* discretization error estimates, and 3) the adaptive meshes proposed in this study. By conventional error estimates, we mean the *a priori* discretization error estimates (3.14) and (3.15) for the solution of the forward

Table 3.4: The L^1 norm of α^λ and the error $\|\alpha^\lambda - \tilde{\alpha}_n^\lambda\|_{L^1(\Omega)}$ for each experiment described in the simulation study 2. The first column shows the type of the meshes used in the forward and inverse problems, respectively. The radius of the circular heterogeneity is given in cm.

	Radius:	0.50	0.75	1.0	1.25
	$\|\alpha^\lambda\|_{L^1(\Omega)}$:	0.7196	1.3760	1.4759	1.7817
Uniform-Uniform	$\|\alpha^\lambda - \tilde{\alpha}_n^\lambda\|_{L^1(\Omega)}$:	0.5622	0.5706	0.5850	0.6337
Adaptive-Uniform	$\|\alpha^\lambda - \tilde{\alpha}_n^\lambda\|_{L^1(\Omega)}$:	0.2153	0.2776	0.3766	0.5113
Adaptive-Adaptive	$\|\alpha^\lambda - \tilde{\alpha}_n^\lambda\|_{L^1(\Omega)}$:	0.2020	0.2630	0.3592	0.5034

problem, and the *a priori* interpolation error estimate (3.11) for the solution of the inverse problem.

To simulate the optical data, we use the same source-detector configuration considered in the first simulation study. We simulate the optical data by solving the diffusion equation at $\omega = 0$ on a fine uniform grid with 61 nodes along x and y directions, where the refractive index mismatch parameter $a = 3$. We note that, in all reconstructions, the background absorption value is set to $\mu_a = 0.04$ cm^{-1} and the diffusion coefficient D is assumed to be constant and $D(\mathbf{x}) = 0.0410$ cm, for all $\mathbf{x} \in \Omega \cup \partial\Omega$.

The uniform meshes used in this simulation study are identical to the ones used in simulation studies 1 and 2. Sample meshes for the forward problem solution using the conventional and the proposed adaptive meshing strategies are shown in Figures 3.7(a)-3.7(b) and Figures 3.8(c) and 3.8(d), respectively. We see that the conventional adaptive mesh generation strategy leads to meshes refined around only sources or detectors, but not both. In contrast, Figures 3.8(c) and 3.8(d) show that the proposed strategy results in adaptive meshes refined around sources, detectors, and the heterogeneity as well. This observation is consistent with Theorem 2. The adaptive mesh for the inverse problem solution, which was generated using the *a priori* interpolation error estimate (3.11) is shown in Figure 3.8(f). Note that the mesh was generated for the case where the circular heterogeneity was centered at $(3.0, 3.5)$. The mesh generated based on Theorem 1 (Figure 3.8(e)) provides higher resolution close to the sources and detectors as compared to the mesh shown in Figure 3.8(f), which is merely refined around the heterogeneity.

In this simulation study, we consider 4 different positions for the center of the circular heterogeneity with radius 0.75 cm, along $y-$axis: center at (3.0,3.0), (3.0,3.5), (3.0,4.0), and (3.0,4.5), respectively. Similar to the previous simulations, we compute the error in the reconstructed images for all cases, and compare the error values attained by different meshing strategies. Finally we present the reconstructed images obtained by using different mesh strategies corresponding to the case where the circular inclusion is centered at $(3.0, 3.5)$ and $(3.0, 4.0)$.

Using the meshes for the forward problem discretization (see Figures 3.7(a)-3.7(b)), which were generated by using the conventional *a priori* discretization error estimates (3.14)-(3.15), leads to the image reconstructions shown in Figures 3.7(c) and 3.7(d), where the regularization parameter $\lambda = 10^{-8}$. We observe that the finite dimensional operator does not provide a stable solution. We note that using an adaptive mesh for the inverse problem solution does not change the outcome (Figure 3.7(d)). Note also that the meshes generated by using the conventional *a priori* discretization error estimates (3.14)-(3.15) are sufficient to provide accurate finite element approximations to the actual solutions g_j and g_i^*. Therefore, the unstable reconstructions can be attributed to the errors $\mathcal{K} - \tilde{\mathcal{K}}$ and $\gamma - \tilde{\gamma}$, due to inappropriate discretization as noted by Theorem 2. In consistence with Theorem 2, this observation suggests that solving the forward problem accurately does not necessarily imply that approximate operator $\tilde{\mathcal{K}}$ and $\tilde{\gamma}$ are error-free. Therefore, in order to address such problems, one has to follow a discretization scheme based on Theorem 2 for the solution of the forward problem, which takes into account the interaction between the solutions of the diffusion equation and the associated adjoint problem, as described in Section 3.3.1.

In order to suppress the severe artifacts observed in Figures 3.7(c) and 3.7(d), we increased the regularization parameter and set $\lambda = 10^{-6}$. The resulting images are shown in Figures 3.7(e) and 3.7(f). As noted by Theorems 1 and 2, increasing the regularization parameter reduces the error in the reconstructed images. However, increasing the regularization parameter will also compromise the image quality and lead to over-smoothed images. In order to address the instability issue without degrading the image quality by using high regularization parameters, we modified the adaptive mesh generation method that lead to the meshes shown in Figures 3.7(a)-3.7(b). In this context, for the first 2 refinements, we used the proposed mesh generation algorithm based on Theorem 2 to generate an initial adaptive mesh; and for the next 2 refinements, we used the conventional error estimates (3.14)-(3.15). Following this modification, the samples of the resulting adaptive meshes are shown in Figures 3.8(a) and 3.8(b). For a comparison, we also present in Figures 3.8(c)-3.8(d), the adaptive meshes generated by using the proposed adaptive mesh generation algorithms as described in Section 3.3.1 and Remark 1(iii). We observe that the meshes shown in Figures 3.8(c)-3.8(d), indicate further refinement around sources, detectors and the circular heterogeneity as compared to the adaptive meshes shown in Figures 3.8(a) and 3.8(b).

Examples of the adaptive meshes generated for the inverse problem based on Theorem 1 and the conventional *a priori* interpolation estimate (3.11) are shown in Figures 3.8(e) and 3.8(f), respectively. We observe that the adaptive mesh shown in Figure 3.8(e) provides higher resolution around sources and detectors as compared to the adaptive mesh shown in Figure 3.8(f).

We note that the uniform meshes used in this simulation study are identical to those used in the previous simulation studies.

In order to compare the performance of the conventional and proposed adaptive mesh strategies, we perform 4 experiments and compute the error in the reconstructed optical absorption images. For each experiment, we consider 5 different mesh strategies and refer to Table 3.5 for the description of these experiments.

We show the reconstructed optical absorption images for the two cases in Figures 3.9 and 3.10, corresponding to the circular heterogeneity centered at (3.0,3.5) and (3.0,4.0), respectively. Figures 3.9(a) and 3.10(a) show the reference absorption image reconstructions which are used to compute the error in the reconstructed optical images.

Figure 3.9(b) shows the image reconstructed using coarse uniform meshes for both the forward and inverse problems, for the case where the circular inclusion is centered at (3.0,3.5) where the regularization parameter was set to $\lambda = 10^{-8}$. With the same value of the regularization parameter, Figure 3.9(c) shows the re-

Table 3.5: The relevant parameters in the experiments 1-5 in simulation study 3. The abbreviation "Conv." implies that the corresponding mesh was generated using the conventional *a priori* discretization error estimates (3.14)-(3.15) for the forward problem solution, and the conventional *a priori* interpolation error estimate (3.11) for the inverse problem solution. The abbreviation "Prop." refers to the adaptive meshes generated by using the proposed adaptive mesh generation algorithms based on Theorems 1 and 2, for the inverse and forward problem solutions, respectively. The last column in the table shows the coordinates of the center of the circular heterogeneity, considered in each experiment.

	Mesh (Forward)	**Mesh (Inverse)**	Center at:
Exp. 1	Uniform	Uniform	[(3.0,3.0), (3.0,3.5), (3.0,4.0), (3.0,4.5)]
Exp. 2	Adaptive (Conv.)	Uniform	[(3.0,3.0), (3.0,3.5), (3.0,4.0), (3.0,4.5)]
Exp. 3	Adaptive (Conv.)	Adaptive (Conv.)	[(3.0,3.0), (3.0,3.5), (3.0,4.0), (3.0,4.5)]
Exp. 4	Adaptive (Prop.)	Uniform	[(3.0,3.0), (3.0,3.5), (3.0,4.0), (3.0,4.5)]
Exp. 5	Adaptive (Prop.)	Adaptive (Prop.)	[(3.0,3.0), (3.0,3.5), (3.0,4.0), (3.0,4.5)]

constructed image by using the adaptive mesh based on Theorem 2 for the forward problem and the coarse uniform mesh (shown in Figure 3.2(b)) for the inverse problem. Figure 3.9(e) shows the reconstructed image obtained by using the adaptive meshes based on Theorems 1 and 2. We observe the improvements especially around the boundaries. Using the conventional adaptive meshes for the forward problem solution, which were modified around sources and detectors as noted before, we ran into a similar instability problem. Therefore, in order to obtain better reconstructions with the conventional adaptive meshes, we set the regularization parameter $\lambda = 10^{-7}$ in the corresponding inverse problem formulations. The resulting reconstructed images are shown in Figures 3.9(d) and 3.9(f). In this case, we observe that the use of conventional adaptive meshes for the forward and inverse problems does not improve the image quality as compared to the reconstructed image shown in Figure 3.9(b), which is obtained by using coarse uniform meshes.

We observe similar results for the case where the circular inclusion is centered at (3.0,4.0). We note that the regularization parameter is set to $\lambda = 10^{-8}$ for all reconstructions except for the reconstructions obtained by using conventional adaptive meshes, in which case $\lambda = 10^{-7}$. Figures 3.9 show the reconstructed images corresponding to all meshing strategies.

Table 3.6 shows the error norm computations for all cases. The error values are consistent with Figures 3.9 and 3.10. In all cases, the proposed adaptive meshes significantly reduce the error in the reconstructed images. Furthermore, the image quality is enhanced by merely appropriate discretization, without having to increase the regularization parameter. In contrast, the conventional adaptive meshes perform worse than uniform meshes even though a higher regularization parameter is used.

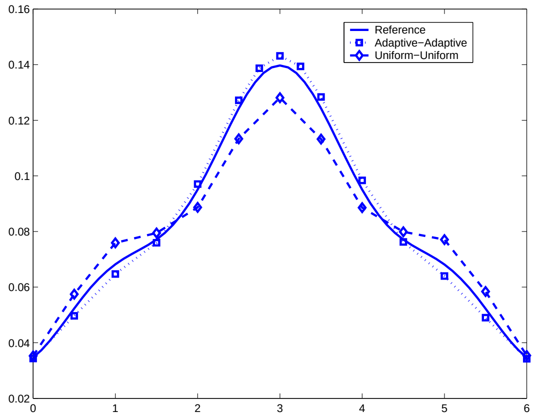
Table 3.6: The error $\|\alpha^\lambda - \tilde{\alpha}_n^\lambda\|_{L^1(\Omega)}$ for each experiment described in the simulation study 3. The first column shows the type of the meshes used in the forward and inverse problems, respectively. The superscript 'C' denotes that the corresponding adaptive mesh generation is based on the conventional *a priori* error estimates (3.14)-(3.15) and (3.11).

	Radius at:	(3.0,3.0)	(3.0,3.5)	(3.0,4.0)	(3.0,4.5)
Uniform-Uniform	$\|\alpha^\lambda - \tilde{\alpha}_n^\lambda\|_{L^1(\Omega)}$:	0.4539	0.4606	0.4733	0.4956
Adaptive-Uniform	$\|\alpha^\lambda - \tilde{\alpha}_n^\lambda\|_{L^1(\Omega)}$:	0.2690	0.2695	0.2634	0.2507
Adaptive-Adaptive	$\|\alpha^\lambda - \tilde{\alpha}_n^\lambda\|_{L^1(\Omega)}$:	0.2433	0.2455	0.2459	0.2434
Adaptive-Uniform	$\|\alpha^\lambda - \tilde{\alpha}_n^{\lambda,C}\|_{L^1(\Omega)}$:	0.7989	0.7596	0.7072	0.6418
Adaptive-Adaptive	$\|\alpha^\lambda - \tilde{\alpha}_n^{\lambda,C}\|_{L^1(\Omega)}$:	0.8011	0.7614	0.7070	0.6351

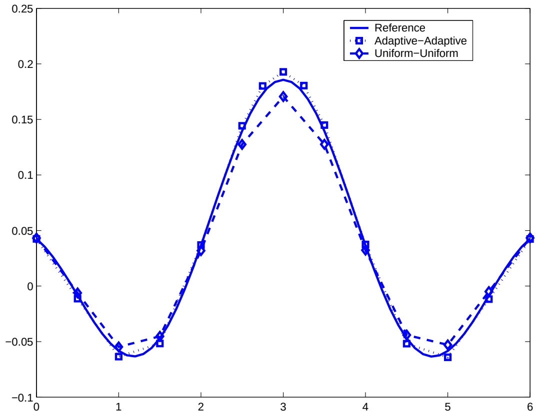
3.5 Conclusion

In this work, based on the error analysis presented in part I [39], we developed two new adaptive mesh generation algorithms, one for the forward and one for the inverse problem, which take into account the interdependence between the solutions of the two problems. We have also presented the computational complexity of the presented adaptive mesh generation algorithms. Our numerical experiments provided a verification of Theorems 1 and 2 and showed that the proposed mesh generation algorithms significantly improve the accuracy of the reconstructed optical images for a given number of unknowns in the discrete forward and inverse problems. Conventional error estimates do not include domain specific factors. As a result, the adaptive mesh generation algorithms based on conventional error estimates (3.14)-(3.15) and (3.11) may lead to high errors in reconstructed optical images as demonstrated in our numerical experiments.

We finally note that the adaptive mesh generation algorithms introduced in this paper can be adapted for the forward and inverse problems of similar inverse parameter estimation problems, such as electrical impedance tomography, optical fluorescence tomography, bioluminescence tomography, microwave imaging etc.

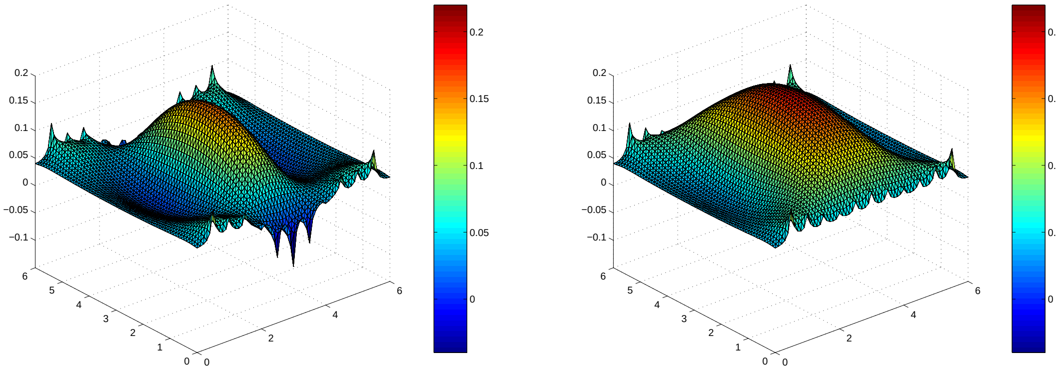


(a) The cross-sectional cuts taken from Figures 3.3(a), 3.3(b), and 3.3(d), along x direction at $y = 3$. The solid, square, and diamond lines correspond to the cross-sectional cuts taken from the images shown in Figures 3.3(a), 3.3(d), and 3.3(b), respectively.



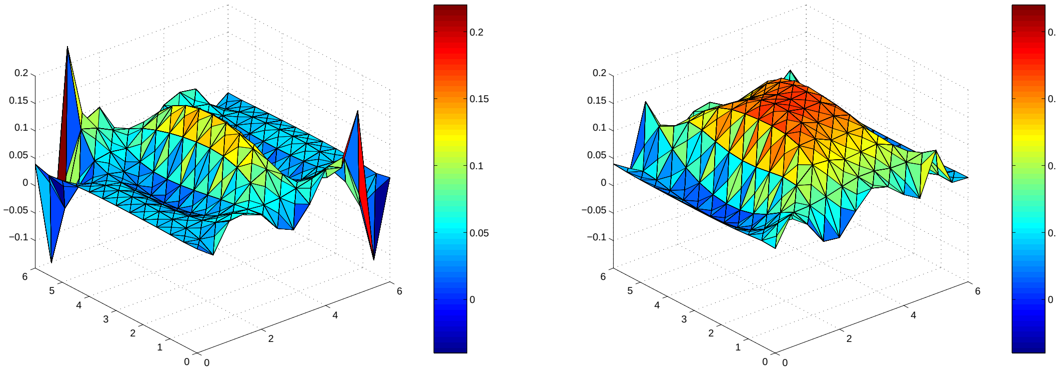
(b) The cross-sectional cuts taken from Figures 3.4(a), 3.4(b), and 3.4(d), along x direction at $y = 3$. The solid, square, and diamond lines correspond to the cross-sectional cuts taken from the images shown in Figures 3.4(a), 3.4(d), and 3.4(b), respectively.

Figure 3.5: The cross-sectional views from the reconstructed images in simulation study 1, corresponding to the cases where the background $\mu_a = 0.032$ and $\mu_a = 0.050$ cm^{-1}, respectively.



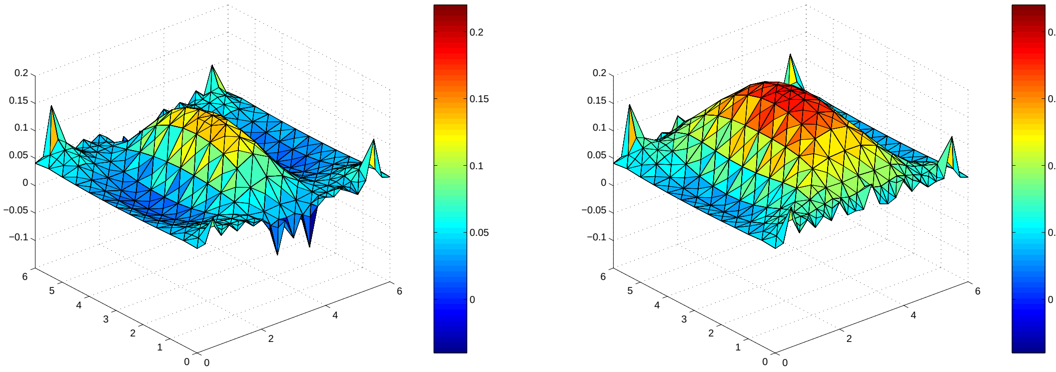
(a) The optical absorption image used as the reference for error computations. The image corresponds to the reconstruction of the circular heterogeneity of radius 0.5 cm.

(b) The optical absorption image used as the reference for error computations. The image corresponds to the reconstruction of the circular heterogeneity of radius 1.25 cm.



(c) The reconstructed absorption image using the uniform mesh in Figure 3.2(a) for the forward, and the uniform mesh in Figure 3.2(b) for the inverse problem.

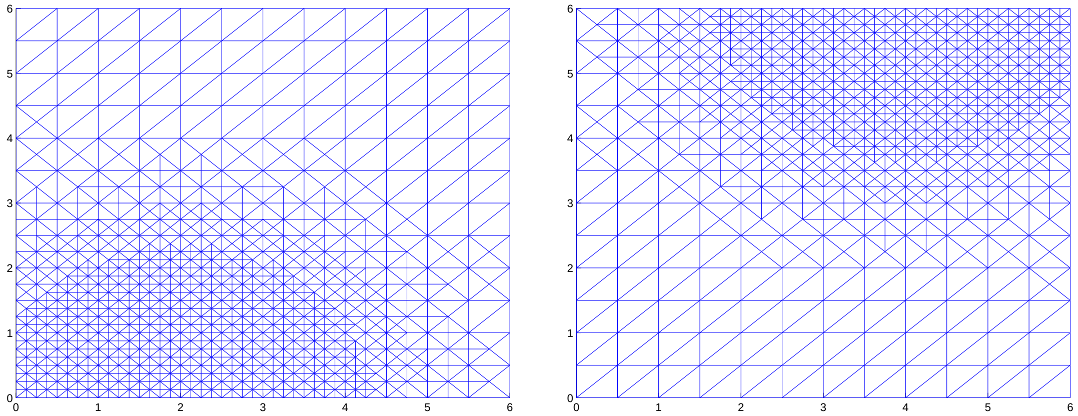
(d) The reconstructed absorption image using the uniform mesh in Figure 3.2(a) for the forward, and the uniform mesh in Figure 3.2(b) for the inverse problem.



(e) The reconstructed absorption image using adaptive meshes for both the forward and the inverse problems.

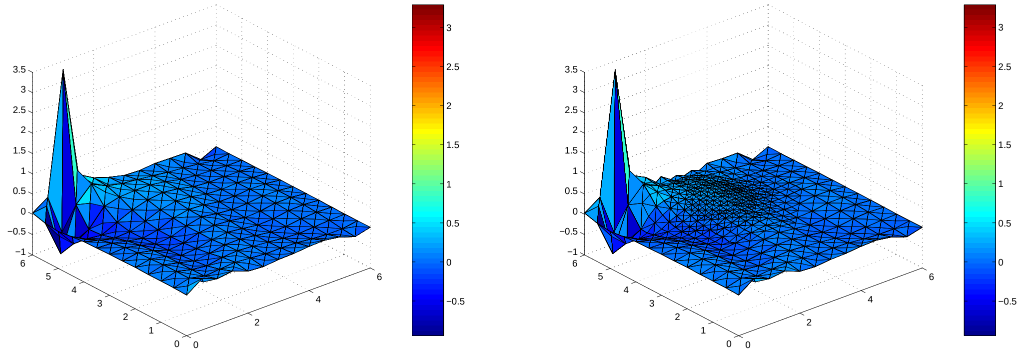
(f) The reconstructed absorption image using an adaptive mesh for both the forward and the inverse problems.

Figure 3.6: The results of simulation study 2. The left and right columns show the reconstructed images regarding the optical heterogeneity with radius 0.50 cm, and 1.25 cm, respectively. The background $\mu_a = 0.040$ cm^{-1} in all of the reconstructions. The reference images shown in (a) and (b) are obtained using a uniform mesh with 61×61 nodes in both the forward and inverse problems.



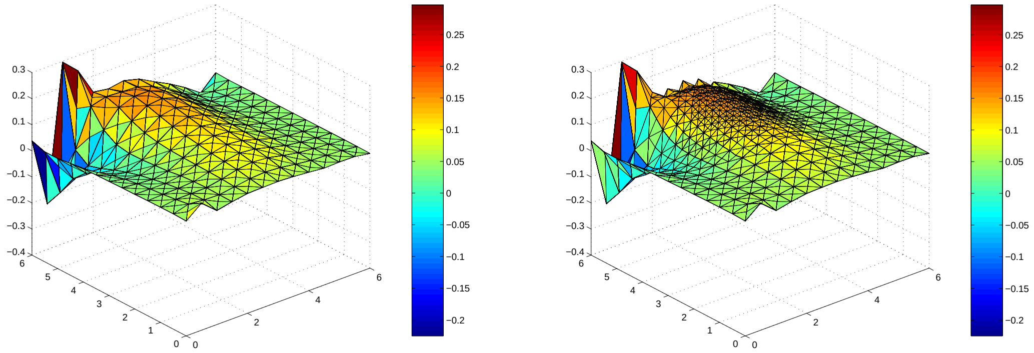
(a) The adaptive mesh with 865 nodes for the forward problem solution for the source located at (2.0,0), generated based on the conventional *a priori* error estimate (3.14).

(b) The adaptive mesh with 865 nodes for the forward problem solution for the detector located at (4.0,6.0), generated based on the conventional *a priori* error estimate (3.15).



(c) The reconstructed optical image using adaptive mesh for the forward and uniform mesh for the inverse problem. $\lambda = 10^{-8}$.

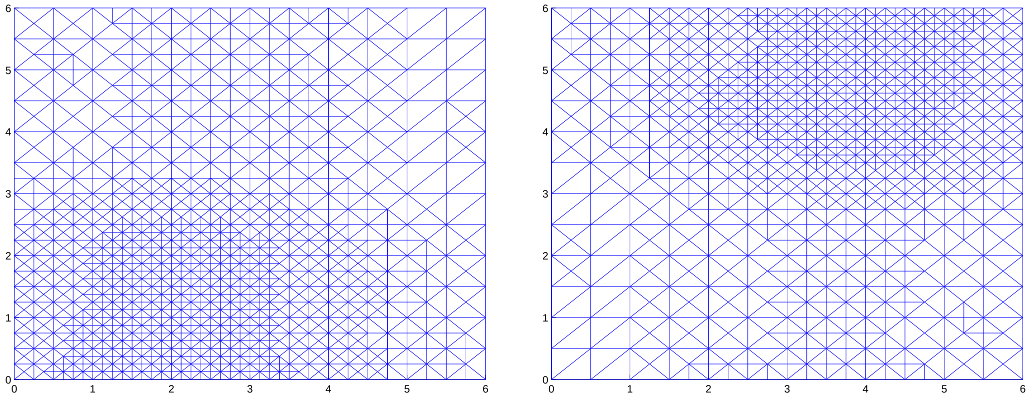
(d) The reconstructed optical image using adaptive meshes for both the forward and inverse problems.$\lambda = 10^{-8}$.



(e) The reconstructed optical image using adaptive mesh for the forward and uniform mesh for the inverse problem. $\lambda = 10^{-6}$.

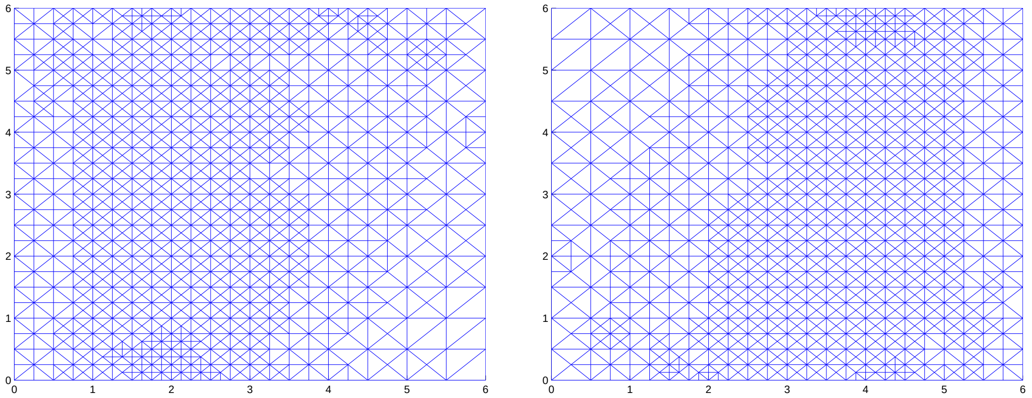
(f) The reconstructed optical image using adaptive meshes for both the forward and inverse problems. $\lambda = 10^{-6}$.

Figure 3.7: (a)-(b) Samples of adaptive meshes in the third simulation study, generated by using the conventional error estimates (3.14) and (3.15), which led to unstable optical image reconstruction shown in (c) to (f), for the circular heterogeneity centered at $(3.0, 3.5)$. (c)-(d) The unstable optical image reconstructions in the third simulation study, obtained by using the adaptive meshes for the forward problem solution whose examples are shown in (a)-(b). $\lambda = 10^{-8}$. (e)-(f) The unstable optical image reconstructions in the third simulation study, obtained by using the adaptive meshes for the forward problem solution whose examples are shown in (a)-(b). λ



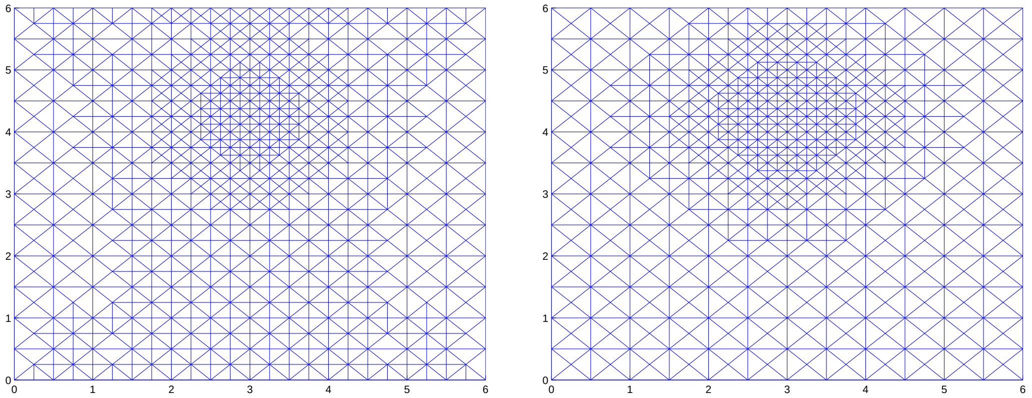
(a) The adaptive mesh with 942 nodes for the forward problem solution for the source located at $(2.0, 0)$, obtained by refining the adaptive mesh shown in Figure 3.7(a) around the detectors.

(b) The adaptive mesh with 955 nodes for the forward problem solution for the detector located at $(4.0, 6.0)$, obtained by refining the adaptive mesh shown in Figure 3.7(b) around the sources.



(c) The adaptive mesh with 895 nodes for the forward problem solution for the source located at $(2.0, 0)$, generated based on Theorem 2.

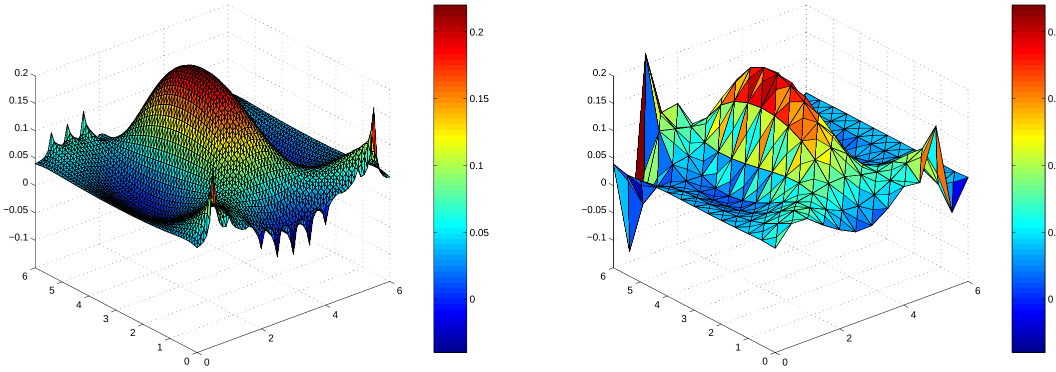
(d) The adaptive mesh with 896 nodes for the forward problem solution for the detector located at $(4.0, 6.0)$, generated based on Theorem 2.



(e) The adaptive mesh with 691 nodes for the inverse problem solution, generated based on Theorem 1.

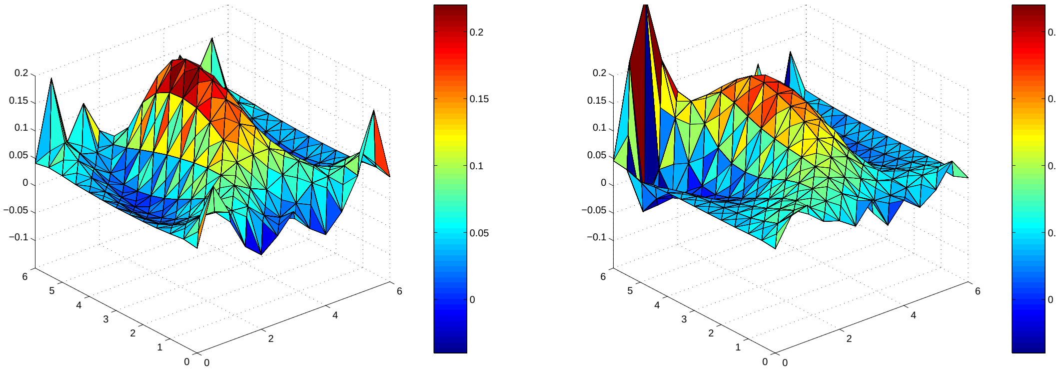
(f) The adaptive mesh with 609 nodes for the inverse problem solution, generated based on the conventional error estimate (3.11).

Figure 3.8: Samples of adaptive meshes used in the third simulation study, which led to the optical image reconstructions shown in Figure 3.10. The meshes were generated for the circular heterogeneity centered at (3.0,4.5).



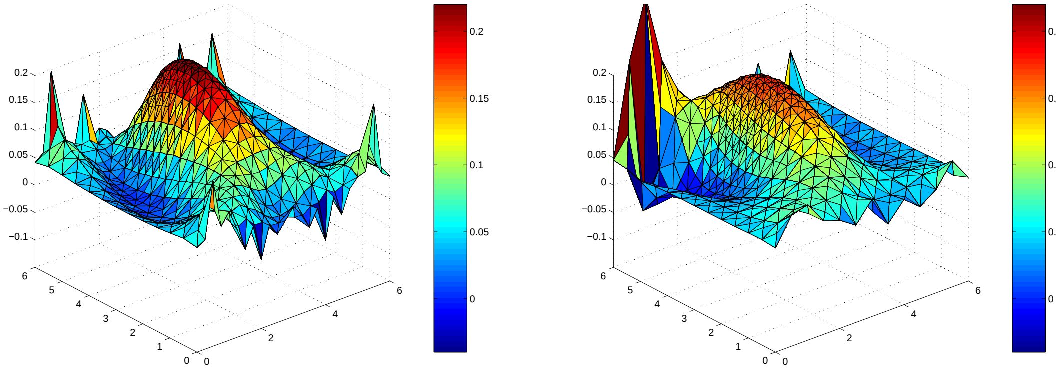
(a) The absorption image used as the reference in the error computations.

(b) The reconstructed absorption image using the uniform mesh in Figure 3.2(a) for the forward, and the uniform mesh in Figure 3.2(b) for the inverse problem.



(c) The reconstructed absorption image using adaptive meshes based on Theorem 2 for the forward, and the uniform mesh in Figure 3.2(b) for the inverse problem.

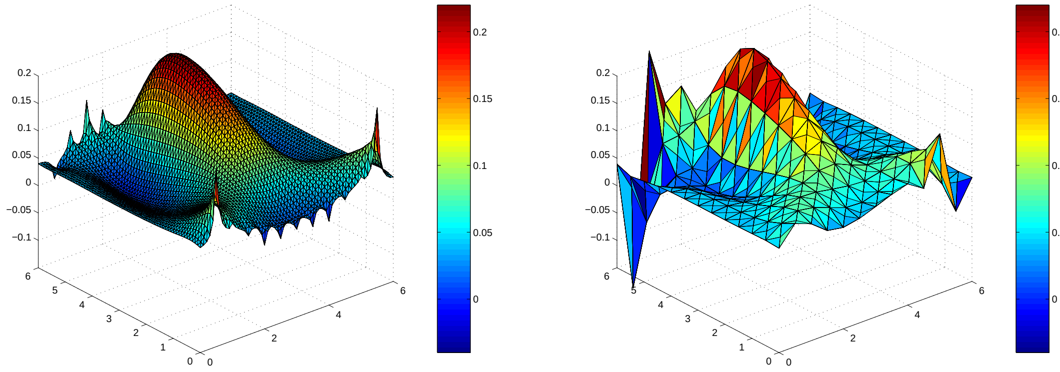
(d) The reconstructed absorption image using adaptive meshes based on *a priori* error estimates (3.14) and (3.15) for the forward, and the uniform mesh in Figure 3.2(b) for the inverse problem.



(e) The reconstructed absorption image using adaptive meshes based on Theorem 2 for the forward, and using the adaptive mesh based on Theorem 1 for the inverse problem.

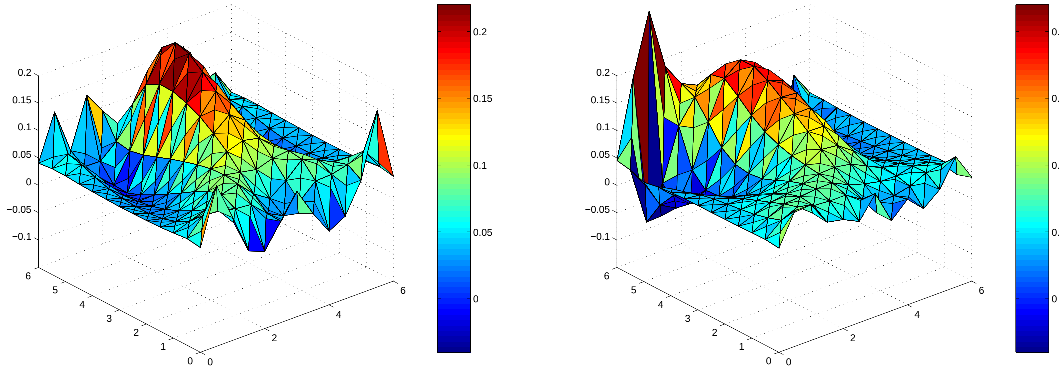
(f) The reconstructed absorption image using adaptive meshes based on *a priori* error estimates (3.14) and (3.15) for the forward, and the interpolation error estimate (3.11) for the inverse problem.

Figure 3.9: The reconstructed optical images regarding the circular heterogeneity centered at $(3.0, 3.5)$ in the third simulation study.



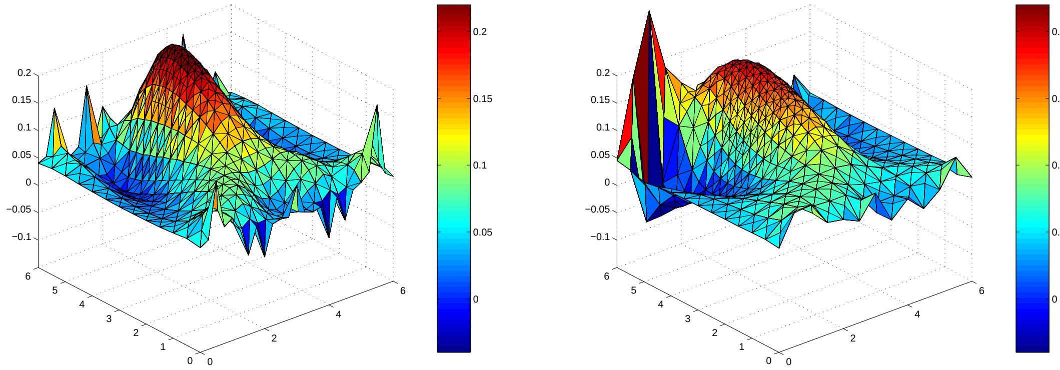
(a) The absorption image used as the reference in the error computations.

(b) The reconstructed absorption image using the uniform mesh in Figure 3.2(a) for the forward, and the uniform mesh in Figure 3.2(b) for the inverse problem.



(c) The reconstructed absorption image using adaptive meshes based on Theorem 2 for the forward, and the uniform mesh in Figure 3.2(b) for the inverse problem.

(d) The reconstructed absorption image using adaptive meshes based on *a priori* error estimates (3.14) and (3.15) for the forward, and the uniform mesh in Figure 3.2(b) for the inverse problem.



(e) The reconstructed absorption image using adaptive meshes based on Theorem 2 for the forward, and the adaptive mesh based on Theorem 1 for the inverse problem.

(f) The reconstructed absorption image using adaptive meshes based on on *a priori* error estimates (3.14) and (3.15) for the forward, and the interpolation error estimate (3.11) for the inverse problem.

Figure 3.10: The reconstructed optical images regarding the circular heterogeneity centered at $(3.0, 4.0)$ in the third simulation study.

CHAPTER 4

Error in Optical Absorption Images due to Born Approximation in Diffuse Optical Tomography

4.1 Introduction

Diffuse Optical Tomography (DOT) poses a challenging nonlinear inverse problem, whose objective is to estimate the unknown optical parameters from boundary measurements [4].

There are a variety of factors that affect the accuracy of DOT imaging, such as model mismatch (due to light propagation model and/or linearization of the inverse problem), measurement noise, discretization, numerical algorithm efficiency, and inverse problem formulation. In this work, we analyze the effect of Born approximation on the accuracy of DOT imaging, in which the light propagation is modelled by the frequency-domain diffusion equation. In this respect, we first derive an upper bound for the norm of the scattered optical field. Then, we show the effect of Born approximation on the accuracy of reconstructed optical absorption images and determine the parameters that control the extent to which the Born approximation affects the imaging accuracy.

The nonlinearity of the associated inverse problem makes DOT imaging a computationally intense problem, which calls for the use of nonlinear inversion methods. In general, nonlinear inversion algorithms require repetitive solution of the forward problem defined by the light propagation model, which is followed by the update of the inverse problem solution. As a result, nonlinear algorithms provide accuracy at the expense of high computational complexity. A widely used approach to overcome the difficulties posed by the nonlinearity is to linearize the inverse problem using Born (or Rytov) approximation and solve the resulting linear problem to reconstruct the optical parameters [7]. Such approximation schemes assume a small perturbation on the background optical coefficients, which are in general assumed to be spatially homogeneous. Spatially homogeneous backgrounds allow for analytical solutions of the forward problem for certain domain geometries [64] and analytical solutions for the linearized inverse problem as well [63]. Hence, linearization of the inverse DOT problem brings several computational advantages and has been used widely in optical tomographic reconstructions, including clinical applications [68, 48, 70, 61, 33, 16, 50, 37]. As a result, Born approximation can be considered as a compromise between accuracy and computational complexity.

The effect of Born approximation on the accuracy of the solutions of parameter estimation problems (in particular of DOT), has been studied in a limited fashion. Natterer has reported an error bound for the Born approximation for the inverse scattering problem of the Helmholtz equation at fixed frequency [67], in which the associated Green's function satisfies the Sommerfeld condition at infinity and the wave number is spatially invariant. However, the dependence of the error on the spatial orientation of the perturbation over the background with respect to the incident wave has not been taken into account. This makes the resulting error bound less stringent for a small-sized domain, which is typical in DOT. In the area of DOT, Boas showed numerically the effect of Born and Rytov approximation on the accuracy of reconstructed optical images [17], where the optical setup consisted of a spherical heterogeneity embedded in an unbounded medium. The study demonstrates the limitation of the linearized algorithms in DOT, but does not provide an analytical analysis for bounded domains and arbitrary optical media with arbitrary source-detector configurations. A similar empirical study has also been reported by Cheng and Boas [24], in which the effect of background optical properties on the imaging accuracy is shown.

In this work, we model the propagation of light by the frequency-domain diffusion equation with Robin boundary conditions. For the inverse problem, we focus on the estimation of the absorption coefficient. We consider the linear integral equation resulting from the linearization of the inverse problem based on Born approximation and use zeroth-order Tikhonov regularization to address the ill-posedness of the resulting integral equation. We first derive a bound for the norm of the scattered optical field, originating from the presence of a perturbation in the absorption coefficient of an arbitrary bounded optical background. Then, we use this error bound to show the effect of Born approximation on the accuracy of the reconstructed optical absorption images. In addition, we show the effect of zeroth-order Tikhonov regularization on the reconstructed optical image. Our analysis identifies several factors which influence the extent to which Born approximation affects the accuracy of the reconstructed images. For example; the inter-dependence of forward and inverse problem solutions, the orientation of the optical heterogeneity with respect to the sources and the detectors, and the number of sources and detectors. The analysis provides an error bound that takes into account the spatial orientation of the heterogeneity, the optical field generated by the light sources and the detector positions. This makes the derived error bound significant especially in optical media with relatively small sizes. The error analysis also provides a measure for the choice of step length in iterative linearization based nonlinear optimization methods, such as trust-region algorithms [60].

The outline of this paper is as follows: Section 2 defines the scattered field and Born approximation. In Section 3, we define the regularized inverse problem with and without Born approximation, respectively. In Section 4, we present an analysis of the error resulting from Born approximation and zeroth-order Tikhonov regularization. In Section 5, we briefly discuss the implications of the derived error. In Section 6, we present numerical experiments which is followed by the conclusions section.

4.2 The Scattered Field and Born Approximation

In this section, we first describe the model for NIR light propagation in a bounded domain, which models the forward problem and define the scattered field due to a perturbation in the absorption coefficient of the medium. Next, we give the Born approximation to the scattered field. Tables 5.1 and 5.2 provide a list of the notation, and the definition of function spaces and norms used throughout the paper, respectively. We note that we use calligraphic letters to denote the operators, e.g. $\mathcal{A}_a$, $\mathcal{I}$, $\mathcal{K}$ etc. and for any operator $\mathcal{F}$ or function f; $\tilde{\mathcal{F}}$, $\tilde{f}$ denote their relevant approximations which will be made clear in the text.

We use the frequency-domain diffusion equation to model the NIR light propagation in a bounded domain $\Omega \subset \mathbb{R}^3$ [4]:

$$-\nabla \cdot D(\mathbf{x})\nabla u^j(\mathbf{x}) + \left(\mu_a(\mathbf{x}) + \frac{i\omega}{c}\right) u^j(\mathbf{x}) = Q^j(\mathbf{x}) \quad \mathbf{x} \in \Omega, \tag{4.1}$$

$$u^j(\mathbf{x}) + 2aD(\mathbf{x})\frac{\partial u^j}{\partial n}(\mathbf{x}) = 0 \quad \mathbf{x} \in \partial\Omega, \tag{4.2}$$

where $D(\mathbf{x})$ is the spatially varying isotropic diffusion coefficient [55] and $\mu_a(\mathbf{x})$ is the spatially varying absorption coefficient at $\mathbf{x}$, c is the speed of the light, a is a parameter governing the internal reflection at the boundary $\partial\Omega$, and Q^j is the point source located at x_s^j, $j = 1, \cdots, N_s$ with modulation frequency ω [77]. In this work, we approximate the point source Q^j in (5.1) by a Gaussian function with sufficiently low variance, whose center is located at $\mathbf{x}_s^j$ [39].

Let $\mu_a(\mathbf{x}) = \mu_{a0}(\mathbf{x}) + \alpha(\mathbf{x})$, for all $\mathbf{x} \in \Omega$, where α is the perturbation over some $\mu_{a0} > 0$. In (5.1)-(5.2), let $u^j = u_0^j + u_s^j$, where u_0^j is the incident field originating from μ_{a0} and u_s^j is the scattered field due to the perturbation α, for $j = 1, \cdots, N_s$. In other words, u_0^j satisfies the following boundary value problem:

$$-\nabla \cdot D(\mathbf{x})\nabla u_0^j(\mathbf{x}) + \left(\mu_{a0}(\mathbf{x}) + \frac{i\omega}{c}\right) u_0^j(\mathbf{x}) = Q^j(\mathbf{x}) \quad \mathbf{x} \in \Omega, \tag{4.3}$$

$$u_0^j(\mathbf{x}) + 2aD(\mathbf{x})\frac{\partial u_0^j}{\partial n}(\mathbf{x}) = 0 \quad \mathbf{x} \in \partial\Omega. \tag{4.4}$$

Then u_s^j satisfies the following boundary value problem:

$$-\nabla \cdot D(\mathbf{x})\nabla u_s^j(\mathbf{x}) + \left(\mu_{a0}(\mathbf{x}) + \frac{i\omega}{c}\right) u_s^j(\mathbf{x}) = -\alpha(\mathbf{x})u^j(\mathbf{x}) \quad \mathbf{x} \in \Omega, \tag{4.5}$$

$$u_s^j(\mathbf{x}) + 2aD(\mathbf{x})\frac{\partial u_s^j}{\partial n}(\mathbf{x}) = 0 \quad \mathbf{x} \in \partial\Omega. \tag{4.6}$$

Making use of the Green's function, the scattered field u_s^j at $\mathbf{x}$ is given by the Lippmann-Schwinger equation [25]:

$$u_s^j(\mathbf{x}) \quad = \quad -\int_{\Omega} \overline{g(\mathbf{x}, \acute{\mathbf{x}})} u^j(\acute{\mathbf{x}})\alpha(\acute{\mathbf{x}})d\acute{\mathbf{x}} \tag{4.7}$$

where $g(\mathbf{x}; \acute{\mathbf{x}})$ is the Green's function of (4.5)-(4.6) at $\mathbf{x}$, due to a point source located at $\acute{\mathbf{x}}$.

We note that u^j depends implicitly on the unknown perturbation α and hence (4.7) is nonlinear in α. The Born approximation $\tilde{u}_s^j$ to u_s^j is obtained by replacing u^j in (4.7) by u_0^j, which is independent of α [4, 25]:

$$\tilde{u}_s^j(\mathbf{x}) \quad = \quad -\int_\Omega \overline{g(\mathbf{x},\acute{\mathbf{x}})} u_0^j(\acute{\mathbf{x}})\alpha(\acute{\mathbf{x}})d\acute{\mathbf{x}}. \tag{4.8}$$

4.3 Inverse Problem

In this section, we explore the inverse problem of recovering the absorption coefficient from boundary measurements. First, we consider the inverse problem formulation based on (4.7) to obtain a baseline for the analysis of the error resulting from linearization by Born approximation. Next, we consider the linearized inverse problem formulation based on Born approximation (4.8).

For practical considerations and to simplify the notation involved in the inverse problem formulation and the following error analysis, we first define the adjoint problem associated with (4.3)-(4.4) [39]:

$$-\nabla \cdot D(\mathbf{x})\nabla g_i^*(\mathbf{x}) + \left(\mu_{a0}(\mathbf{x}) - \frac{i\omega}{c}\right) g_i^*(\mathbf{x}) = 0 \quad \mathbf{x} \in \Omega, \tag{4.9}$$

$$g_i^*(\mathbf{x}) + 2aD(\mathbf{x})\frac{\partial \phi}{\partial n}(\mathbf{x}) = Q_i^*(\mathbf{x}) \quad \mathbf{x} \in \partial\Omega, \tag{4.10}$$

where Q_i^* denotes the ith adjoint source and g_i^* is the solution of the adjoint problem for the ith adjoint source. For a point adjoint source located at the detector position $\mathbf{x}_d^i$, the following holds [4]:

$$g_i^*(\mathbf{x}) = g(\mathbf{x}_d^i, \mathbf{x}), \quad \mathbf{x} \in \Omega, \tag{4.11}$$

for $i = 1, \cdots, N_d$ where N_d is the number of detectors. In this work, we model the adjoint source by a Gaussian function with sufficiently low variance, centered at $\mathbf{x}_d^i$. Then, $g_i^*(\mathbf{x}) \approx g(\mathbf{x}_d^i, \mathbf{x})$ models the finite size of the detectors and accounts for the smoothing on the detector readings [39].

4.3.1 Inverse problem statement without Born approximation

In order to assess the accuracy of the reconstructed optical absorption images resulting from linearization by Born approximation, we consider (4.7) as the baseline to construct the inverse problem formulation which possesses no error resulting from Born approximation. In this respect, we assume that u^j is *known* independently of the *unknown* α. As a result, (4.7) becomes linear in α given u^j on Ω for $j = 1, \cdots, N_s$. We note that this assumption allows us to employ a linear analysis method to derive an explicit bound for the error in the reconstructed optical absorption images (see Section 4.4).

Consider the integral equation (4.7) and suppose that u^j for $j = 1, \cdots, N_s$ are known. Let $\Gamma_{i,j}$ be the (i,j)th entry in the vector $\Gamma \in \mathbb{C}^{N_d \times N_s}$, which represents the differential measurement at the ith detector due to the jth source such that $\Gamma_{i,j} := u_s^j(\mathbf{x}_d^i)$, where $\mathbf{x}_d^i$ denotes the position of the ith detector, $i = 1, \cdots, N_d$. Then, the following linear equation relates the differential boundary measurements to α [4, 39]:

$$\begin{aligned} \Gamma_{i,j} &= -\int_\Omega \overline{g_i^*(\mathbf{x})} u^j(\mathbf{x}) \alpha(\mathbf{x}) d\mathbf{x} && (4.12) \\ &:= \int_\Omega H_{i,j}(\acute{\mathbf{x}}) \alpha(\acute{\mathbf{x}}) d\acute{\mathbf{x}} \\ &:= (\mathcal{A}_a \alpha)_{i,j}, && (4.13) \end{aligned}$$

where $H_{i,j} := -\overline{g_i^*} u^j$ is the (i,j)th kernel of the matrix-valued operator $\mathcal{A}_a : L^2(\Omega) \rightarrow \mathbb{C}^{N_d \times N_s}$, g_i^* is weak the solution of (4.9)-(4.10), and u^j is the weak solution of (5.1)-(5.2) [39].

The linear operator $\mathcal{A}_a : L^2(\Omega) \rightarrow \mathbb{C}^{N_d \times N_s}$ defined by (5.9) is compact [39]. Hence, (5.9) is ill-posed for the given solution space $L^2(\Omega)$ for α. Regularizing (5.9) with a zeroth-order Tikhonov regularization yields the following equation:

$$\begin{aligned} \gamma = \mathcal{A}_a^* \Gamma &= (\mathcal{A}_a^* \mathcal{A}_a + \lambda \mathcal{I}) \alpha^\lambda && (4.14) \\ &:= \mathcal{K} \alpha^\lambda. && (4.15) \end{aligned}$$

In this representation, $\mathcal{I}$ is the identity operator and $\mathcal{A}_a^* : \mathbb{C}^{N_d \times N_s} \rightarrow L^2(\Omega)$ is the adjoint of $\mathcal{A}_a$, defined by

$$(\mathcal{A}_a^*\beta)(\mathbf{x}) = \sum_{i,j}^{N_d,N_s} H_{i,j}^*(\mathbf{x})\beta_{i,j} = \sum_{i,j}^{N_d,N_s} -g_i^*(\mathbf{x})\overline{g_j(\mathbf{x})}\beta_{i,j}, \tag{4.16}$$

for all $\beta \in \mathbb{C}^{N_d \times N_s}$, where $H_{i,j}^* := -g_i^*\overline{g_j}$ is the (i,j)th kernel of the adjoint operator. Let $\mathcal{A} := \mathcal{A}_a^*\mathcal{A}_a$, then the linear integral operator $\mathcal{A} : L^2(\Omega) \to L^2(\Omega)$ is defined as follows:

$$\begin{aligned}(\mathcal{A}\alpha)(\mathbf{x}) &= \textstyle\sum_{i,j}^{N_d,N_s} H_{i,j}^*(\mathbf{x}) \int_\Omega H_{i,j}(\acute{\mathbf{x}})\alpha(\acute{\mathbf{x}})d\acute{\mathbf{x}} \\ &:= \textstyle\int_\Omega \kappa(\mathbf{x},\acute{\mathbf{x}})\alpha(\acute{\mathbf{x}})d\acute{\mathbf{x}},\end{aligned} \tag{4.17}$$

where $\kappa(\mathbf{x},\acute{\mathbf{x}})$ stands for the kernel of the integral operator $\mathcal{A}$ and is given by

$$\kappa(\mathbf{x},\acute{\mathbf{x}}) = \sum_{i,j}^{N_d,N_s} H_{i,j}^*(\mathbf{x})H_{i,j}(\acute{\mathbf{x}}). \tag{4.18}$$

4.3.2 Inverse problem based on Born approximation

Replacing u^j in (5.9) by u_0^j, the inverse optical absorption problem linearized by Born approximation reads

$$\begin{aligned}\Gamma_{i,j} &= -\int_\Omega \overline{g_i^*(\mathbf{x})}u_0^j(\acute{\mathbf{x}})\tilde{\alpha}(\acute{\mathbf{x}})d\acute{\mathbf{x}} && (4.19)\\ &:= \int_\Omega \tilde{H}_{i,j}(\acute{\mathbf{x}})\tilde{\alpha}(\acute{\mathbf{x}})d\acute{\mathbf{x}} \\ &:= (\tilde{A}_a\tilde{\alpha})_{i,j}, && (4.20)\end{aligned}$$

where the linear matrix-valued operator $\tilde{\mathcal{A}}_a : L^2(\Omega) \to \mathbb{C}^{N_d \times N_s}$ is an approximation to the operator $\mathcal{A}_a$ as a result of Born approximation, $\tilde{H}_{i,j}(\acute{\mathbf{x}}) = \overline{g_i^*(\mathbf{x})}u_0^j(\acute{\mathbf{x}})$ is the (i,j)th kernel of $\tilde{\mathcal{A}}_a$ at $\acute{\mathbf{x}}$, and $\tilde{\alpha} \in L^2(\Omega)$ is an approximation to the actual solution α, as a result of the linearization by Born approximation.

Similar to the operator $\mathcal{A}_a$, the linear operator $\tilde{A}_a : L^2(\Omega) \to \mathbb{C}^{N_d \times N_s}$ defined by (4.19) is compact and this implies the ill-posedness of (4.19). Hence, we regularize (4.19) with a zeroth-order Tikhonov regularization:

$$\begin{aligned} \tilde{\gamma} := \tilde{\mathcal{A}}_a^*\Gamma &= \left(\tilde{\mathcal{A}}_a^*\tilde{\mathcal{A}}_a + \lambda\mathcal{I}\right)\tilde{\alpha}^\lambda \qquad (4.21) \\ &:= \tilde{\mathcal{K}}\tilde{\alpha}^\lambda, \qquad (4.22) \end{aligned}$$

where $\lambda > 0$ is the regularization parameter and $\tilde{\alpha}^\lambda$ is an approximation to $\tilde{\alpha}$. In this representation, $\mathcal{I}$ is the identity operator and $\tilde{\mathcal{A}}_a^* : \mathbb{C}^{N_d \times N_s} \to L^2(\Omega)$ is the adjoint of $\tilde{\mathcal{A}}_a$, defined similar to $\mathcal{A}_a^*$ by the kernel $\tilde{H}_{i,j}^* = \overline{g_i^*}u_0^j$.

Then, the linear integral operator $\tilde{\mathcal{A}} := \tilde{\mathcal{A}}_a^*\tilde{\mathcal{A}}_a : L^2(\Omega) \to L^2(\Omega)$ is defined as follows:

$$\begin{aligned} (\tilde{\mathcal{A}}\tilde{\alpha})(\mathbf{x}) &= \sum_{i,j}^{N_d,N_s} \tilde{H}_{i,j}^*(\mathbf{x}) \int_\Omega \tilde{H}_{i,j}(\acute{\mathbf{x}})\tilde{\alpha}(\acute{\mathbf{x}})d\acute{\mathbf{x}} \\ &:= \int_\Omega \tilde{\kappa}(\mathbf{x}, \acute{\mathbf{x}})\tilde{\alpha}(\acute{\mathbf{x}})d\acute{\mathbf{x}}, \qquad (4.23) \end{aligned}$$

where $\tilde{\kappa}(\mathbf{x}, \acute{\mathbf{x}})$ is the kernel of $\tilde{\mathcal{A}}$ and is given by

$$\tilde{\kappa}(\mathbf{x}; \acute{\mathbf{x}}) = \sum_{i,j}^{N_d,N_s} \tilde{H}_{i,j}^*(\mathbf{x})\tilde{H}_{i,j}(\acute{\mathbf{x}}). \tag{4.24}$$

Comparing $\mathcal{K}$ in (5.19) and $\tilde{\mathcal{K}}$ in (4.22) shows that Born approximation leads to an error in the operator $\tilde{\mathcal{K}}$ with respect to the operator $\mathcal{K}$. This is due to the errors in the kernels $\tilde{H}_{i,j}$ and $\tilde{H}_{i,j}^*$ with respect to $H_{i,j}$ and $H_{i,j}^*$, given by

$$\begin{aligned} \left(H_{i,j} - \tilde{H}_{i,j}\right)(\mathbf{x}) &= -\overline{g_i^*(\mathbf{x})}(u^j - u_0^j)(\mathbf{x}) = \overline{g_i^*(\mathbf{x})}u_s^j(\mathbf{x}) \qquad (4.25) \\ \left(H_{i,j}^* - \tilde{H}_{i,j}^*\right)(\mathbf{x}) &= -g_i^*(\mathbf{x})\overline{(u^j - u_0^j)(\mathbf{x})} = g_i^*(\mathbf{x})\overline{u_s^j(\mathbf{x})}. \qquad (4.26) \end{aligned}$$

As a result, the solution $\tilde{\alpha}^\lambda$ of (4.22) with respect to α^λ in (5.19) has an error which can be attributed to Born approximation. In addition, the zeroth-order Tikhonov regularization implies an error term in the solution α^λ of (5.19) with respect to the absorptive perturbation α. In the following section, we will analyze both errors in the solution $\tilde{\alpha}^\lambda$ of (4.22) with respect to the true absorptive perturbation α.

4.4 Analysis of the error resulting from Born approximation and Tikhonov regularization

In this section, we analyze the effect of Born approximation and Tikhonov regularization on the accuracy of DOT imaging. The analysis is carried out based on the linear inverse problem defined in Section 4.3.1. In this respect, we first derive an upper bound for the norm of the scattered field, which is used to compute the error in the approximations $\tilde{\mathcal{K}}$ and $\tilde{\gamma}$ with respect to $\mathcal{K}$ and γ, respectively. Then, we show the effect of Born approximation on the accuracy of the reconstructed optical image $\tilde{\alpha}^\lambda$ in (4.22) with respect to the solution α^λ of (5.19). Finally, we discuss the effect of Tikhonov regularization on the reconstructed optical image and derive a bound for the error in the reconstructed optical image $\tilde{\alpha}^\lambda$ with respect to the true absorptive perturbation α.

The error in the solution $\tilde{\alpha}^\lambda$ of (4.22) with respect to the actual solution α of (4.13) has two contributors: The error $\tilde{e} = \alpha^\lambda - \tilde{\alpha}^\lambda$, which is the result of Born approximation and the error $e^\lambda = \alpha^\lambda - \alpha$ resulting from Tikhonov regularization. Thus,

$$\tilde{\alpha}^\lambda = \alpha^\lambda - \tilde{e} = \alpha - e^\lambda - \tilde{e}, \tag{4.27}$$

and a bound on the error $\alpha - \tilde{\alpha}^\lambda$ is given by:

$$\|\alpha - \tilde{\alpha}^\lambda\| = \|\tilde{e} + e^\lambda\| \leq \|\tilde{e}\| + \|e^\lambda\|. \tag{4.28}$$

In the following, we will analyze both error contributors to derive the bounds for $\|\tilde{e}\|$ and $\|e^\lambda\|$, respectively, with an emphasis on the norm $\|\tilde{e}\|$ of the error $\tilde{e}$ resulting from Born approximation.

4.4.1 Effect of Born approximation

Before we derive a bound for the error $\tilde{e} = \alpha^\lambda - \tilde{\alpha}^\lambda$, we give the following lemma which provides a bound for the $L^2(\Omega)$ norm of the scattered field u_s^j.

Lemma 1:

The $L^2(\Omega)$ norm of the scattered field u_s^j is bounded by

$$\|u_s^j\|_0 \leq C_s \|u^j \alpha\|_0, \tag{4.29}$$

where C_s is a positive constant.

Proof. By (4.7), the square of the norm $\|u_s^j\|_0$ is given by

$$\begin{aligned} \|u_s^j\|_0^2 &= \int_\Omega d\mathbf{x} \left| \int_\Omega \overline{g(\mathbf{x}; \acute{\mathbf{x}})} u^j(\acute{\mathbf{x}}) \alpha(\acute{\mathbf{x}}) d\acute{\mathbf{x}} \right|^2 \\ &\leq \int_\Omega d\mathbf{x} \left(\int_\Omega |g(\mathbf{x}; \acute{\mathbf{x}}) u^j(\acute{\mathbf{x}}) \alpha(\acute{\mathbf{x}})| d\acute{\mathbf{x}} \right)^2 . \end{aligned} \tag{4.30}$$

Note that $u^j \alpha \in L^2(\Omega)$, which results from the boundedness of α and $L^2(\Omega)$ boundedness of u^j [39]. Then by [59],

$$\|u_s^j\|_0^2 \leq C_s^2 \|u^j \alpha\|_0^2,$$

for some $C_s > 0$. ☐

Note that $u^j = u_0^j + u_s^j$. In order to obtain an explicit bound for $\|u_s^j\|_0$ in terms of u_0^j, (4.29) can be modified as follows:

Corollary 1: Let $\epsilon := C_s \|\alpha\|_\infty < 1$, where C_s is the constant in (4.29). Then,

$$\|u_s^j\|_0 \leq \frac{C_s}{1-\epsilon} \|u_0^j \alpha\|_0. \tag{4.31}$$

Proof. Noting $u^j = u_0^j + u_s^j$, (4.29) can be expanded as $\|u_s^j\|_0 \leq C_s \|u_0^j \alpha\|_0 + C_s \|u_s^j\|_0 \|\alpha\|_\infty$. Letting $\epsilon = C_s \|\alpha\|_\infty$ yields the bound in (4.31). ☐

Remarks:

1. The strength of the scattered field u_s^j depends on the strength of the source Q^j. Hence, $\|u_s^j\|_0$ depends on the strength of the source Q^j.

2. For each source $j = 1, \cdots, N_s$, the scattered field depends on both the field u_0^j and the perturbation α. For a point source, u_0^j will be large close to the source.

Hence, if α is nonzero close to the source, the $L^2(\Omega)$ norm of the scattered field will be large as well.

3. In case Q^j is not a point source, one sufficient (but necessary) condition for (4.29) to hold is $Q^j \in H^1(\Omega)$.

We now analyze the effect of Born approximation on the accuracy of the reconstructed optical image $\tilde{\alpha}^\lambda$. By Theorem 10.1 in [57], the error $\tilde{e}$ in the solution $\tilde{\alpha}^\lambda$ of (4.22) with respect to the solution α^λ of (5.19) is bounded by

$$\|\tilde{e}\| = \|\alpha^\lambda - \tilde{\alpha}^\lambda\| \leq \|\tilde{\mathcal{K}}^{-1}\| \left\{ \|(\tilde{\mathcal{K}} - \mathcal{K})\alpha^\lambda\| + \|\tilde{\gamma} - \gamma\| \right\}. \tag{4.32}$$

Note that the inverse operator $\tilde{\mathcal{K}}^{-1} : L^2(\Omega) \rightarrow L^2(\Omega)$ exists since $\tilde{\mathcal{K}}$ is positive definite for $\lambda > 0$. Furthermore, the inverse operator $\tilde{\mathcal{K}}^{-1}$ is bounded by

$$\|\tilde{\mathcal{K}}^{-1}\|_{L^2(\Omega)\rightarrow L^2(\Omega)} \leq \frac{1}{\lambda}. \tag{4.33}$$

Note that a similar bound can be obtained for $\|\mathcal{K}^{-1}\|_{L^2(\Omega)\rightarrow L^2(\Omega)}$

Lemma 2:

Let $\|g_i^* u_s^j\|_0 \leq \|g_i^* u_0^j\|_0$ for all $j = 1, \cdots, N_s$, $i = 1, \cdots, N_d$ and assume that α and α^λ are bounded. Then, a bound for the error $\tilde{e}$ in the solution $\tilde{\alpha}^\lambda$ of (4.22) with respect to the solution α^λ of (5.19), resulting from Born approximation can be given by

$$\|\alpha^\lambda - \tilde{\alpha}^\lambda\|_0 \leq \frac{C_b}{\lambda} \sum_{i,j}^{N_d,N_s} \left(4\|g_i^* \tilde{\alpha}^\lambda\|_0 + \|\alpha\|_0 \|g_i^*\|_{L^\infty(\Omega)}\right) \|u_0^j \alpha\|_0, \tag{4.34}$$

where C_b is the positive constant given by

$$C_b = \frac{2C_s}{1-\epsilon} \max_{i,j} \|g_i^* u_0^j\|_0. \tag{4.35}$$

Proof. $\|(\tilde{\mathcal{K}} - \mathcal{K})\alpha^\lambda\|_0$ is bounded by [39]

$$\|(\tilde{\mathcal{K}} - \mathcal{K})\alpha^\lambda\|_0 \approx 2\|\mathcal{A}_a^*(\mathcal{A}_a - \tilde{\mathcal{A}}_a)\alpha^\lambda\|_0$$

$$\approx 2 \left\| \sum_{i,j}^{N_d,N_s} g_i^*(\cdot)\overline{u^j(\cdot)} \int_\Omega \left| \left(u^j(\acute{\mathbf{x}})\overline{e_i^*(\acute{\mathbf{x}})} + \overline{g_i^*(\acute{\mathbf{x}})} e_j(\acute{\mathbf{x}}) \right) \alpha^\lambda(\acute{\mathbf{x}}) \right| d\acute{\mathbf{x}} \right\|_0$$
$$\leq 2 \max_{i,j} \|g_i^* u^j\|_0 \sum_{i,j}^{N_d,N_s} \int_\Omega \left| \left(u^j(\acute{\mathbf{x}})\overline{e_i^*(\acute{\mathbf{x}})} + \overline{g_i^*(\acute{\mathbf{x}})} e_j(\acute{\mathbf{x}}) \right) \alpha^\lambda(\acute{\mathbf{x}}) \right| d\acute{\mathbf{x}},$$

where e_i^* and e_j are the errors in the approximations of g_i^* and u^j, respectively. Born approximation results in error for only u^j. Thus, $e_i^* = 0$ and $e_j(\acute{\mathbf{x}}) = (u^j - u_0^j)(\acute{\mathbf{x}}) = u_s^j(\acute{\mathbf{x}})$ for all $\acute{\mathbf{x}} \in \Omega$. Thus, we can write

$$\|(\tilde{\mathcal{K}} - \mathcal{K})\alpha^\lambda\|_0 \leq 2 \max_{i,j} \|g_i^* u^j\|_0 \sum_{i,j}^{N_d,N_s} \|g_i^* \alpha^\lambda\|_0 \|u_s^j\|_0.$$

Noting $\|g_i^* u_s^j\|_0 \leq \|g_i^* u_0^j\|_0$ and $u^j = u_0^j + u_s^j$, above inequality can be expanded as follows:

$$\|(\tilde{\mathcal{K}} - \mathcal{K})\alpha^\lambda\|_0 \leq 2 \max_{i,j} \|g_i^* (u_0^j + u_s^j)\|_0 \sum_{i,j}^{N_d,N_s} \|g_i^* \alpha^\lambda\|_0 \|u_s^j\|_0$$
$$\leq 4 \max_{i,j} \|g_i^* u_0^j\|_0 \sum_{i,j}^{N_d,N_s} \|g_i^* \alpha^\lambda\|_0 \|u_s^j\|_0. \tag{4.36}$$

Note that $\alpha^\lambda = \tilde{\alpha}^\lambda + \tilde{e}$, where $\tilde{e}$ is the error in the reconstructed optical absorption perturbation due to Born approximation. Consistent with the initial assumption, namely $\|g_i^* u_s^j\|_0 \leq \|g_i^* u_0^j\|_0$, we assume

$$\|g_i^* \tilde{e}\|_0 \leq \|g_i^* \tilde{\alpha}^\lambda\|_0,$$

for $i = 1, \cdots, N_d$. Then, (4.36) can be rewritten as follows:

$$\|(\tilde{\mathcal{K}} - \mathcal{K})\alpha^\lambda\|_0 \leq 8 \max_{i,j} \|g_i^* u_0^j\|_0 \sum_{i,j}^{N_d,N_s} \|g_i^* \tilde{\alpha}^\lambda\|_0 \|u_s^j\|_0. \tag{4.37}$$

We first note that for bounded positive D and μ_{a0}, the boundedness of the adjoint source implies g_i^* is bounded [27], thus $g_i^* \in L^\infty(\Omega)$. Following this note, an upper bound for $\|\tilde{\gamma}-\gamma\|_0$ can be obtained as follows [39]:

$$\begin{aligned}
\|\tilde{\gamma}-\gamma\|_0 \approx \left\| \sum_{i,j}^{N_d,N_s} (e_i^*(\cdot)\overline{u^j(\cdot)} + g_i^*(\cdot)\overline{e_j(\cdot)})\Gamma_{i,j} \right\|_0 \\
\leq \max_{i,j} |\Gamma_{i,j}| \sum_{i,j}^{N_d,N_s} \|g_i^* u_s^j\|_0 \\
\leq \max_{i,j} |\Gamma_{i,j}| \sum_{i,j}^{N_d,N_s} \|g_i^*\|^2_{L^\infty(\Omega)} \|u_s^j\|_0, \qquad (4.38)
\end{aligned}$$

since $e_i^* = 0$. A bound for $\max_{i,j}|\Gamma_{i,j}|$ can be given by

$$\max_{i,j} |\Gamma_{i,j}| \leq \max_{i,j} \|g_i^* u^j\|_0 \|\alpha\|_0. \qquad (4.39)$$

Then,

$$\|\tilde{\gamma}-\gamma\|_0 \leq \max_{i,j} \|g_i^* u^j\|_0 \|\alpha\|_0 \sum_{i,j}^{N_d,N_s} \|g_i^*\|^2_{L^\infty(\Omega)} \|u_s^j\|_0, \qquad (4.40)$$

Noting $u^j = u_0^j + u_s^j$ and $\|g_i^* u_s^j\|_0 \leq \|g_i^* u_0^j\|_0$, we write

$$\|\tilde{\gamma}-\gamma\|_0 \leq 2\max_{i,j} \|g_i^* u_0^j\|_0 \|\alpha\|_0 \sum_{i,j}^{N_d,N_s} \|g_i^*\|_{L^\infty(\Omega)} \|u_s^j\|_0. \qquad (4.41)$$

Using (4.31) in (4.37) and (4.41), we set

$$C_b = \frac{2C_s}{1-\epsilon} \max_{i,j} \|g_i^* u_0^j\|_0.$$

Then (4.37) and (4.41) read, respectively

$$\begin{aligned}
\|(\tilde{\mathcal{K}}-\mathcal{K})\alpha^\lambda\|_0 &\leq 4C_b \sum_{i,j}^{N_d,N_s} \|g_i^* \tilde{\alpha}^\lambda\|_0 \|u_0^j \alpha\|_0, \qquad (4.42) \\
\|\tilde{\gamma}-\gamma\|_0 &\leq C_b \|\alpha\|_0 \sum_{i,j}^{N_d,N_s} \|g_i^*\|_{L^\infty(\Omega)} \|u_0^j \alpha\|_0. \qquad (4.43)
\end{aligned}$$

We now use (4.32), (4.42) and (4.43) to obtain (4.34). ☐

Remarks:

1. If $\|u_0^j\alpha\|_0$ is large, then the error in the reconstructed absorption image will be high. This suggests that the support of the heterogeneity α should be away from the sources, where u_0^j for $j = 1, \cdots, N_s$ tends to be large.
2. If the support of the heterogeneity α is close to the detector, then the effect of Born approximation on the overall error is amplified, since this increases the summation term in (4.34).
3. In order to minimize the error $\|\alpha^\lambda - \tilde{\alpha}^\lambda\|$, the support of the heterogeneity must be away from both the sources and detectors.
4. In order to reduce the error $\|\alpha^\lambda - \tilde{\alpha}^\lambda\|$, the sources and detectors which will be used in the reconstruction can be selected based on the support of the heterogeneity α.
5. If there is *a priori* information about the position of the heterogeneity α, the sources and detectors can be located such that the error due to Born approximation is reduced.
6. Lemma 2 is valid for all bounded sources $Q^j \in H^1(\Omega)$.

4.4.2 Effect of Tikhonov regularization

The regularization scheme approximates the solution α of (4.13) by the regularized solution α^λ of (5.19). The following lemma states a bound for the error $e^\lambda = \alpha - \alpha^\lambda$ resulting from zeroth-order Tikhonov regularization.

Lemma 3:

Let $\|g_i^* u_s^j\|_0 \leq \|g_i^* u_0^j\|_0$ for $j = 1, \cdots, N_s$, $i = 1, \cdots, N_d$ and $\alpha \in L^\infty$. Then, a bound for the $L^2(\Omega)$ norm of the error $e^\lambda = \alpha - \alpha^\lambda$, resulting from zeroth-order Tikhonov regularization can be given by

$$\|\alpha - \alpha^\lambda\|_0 \leq C_T \|\alpha\|_0, \tag{4.44}$$

where $C_T > 0$ is given by

$$C_T := 1 + \frac{4}{\lambda} \max_{i,j} \|g_i^* u_0^j\|_0 \tag{4.45}$$

Proof. Clearly,

$$\alpha - \alpha^\lambda = \alpha - \mathcal{K}^{-1}\mathcal{A}\alpha = \alpha - \mathcal{K}^{-1}\mathcal{A}_a^*\Gamma \tag{4.46}$$

and

$$\|\alpha - \alpha^\lambda\|_0 \le \|\alpha\|_0 + \|\mathcal{K}^{-1}\mathcal{A}_a^*\Gamma\|_0.$$

Note that

$$\begin{aligned} \|\mathcal{K}^{-1}\mathcal{A}\Gamma\|_0 &= \left\|\mathcal{K}^{-1}\sum_{i,j}^{N_d,N_s} g_i^*\overline{u^j}\Gamma_{i,j}\right\|_0 \\ &\le \max_{i,j}|\Gamma_{i,j}| \left\|\mathcal{K}^{-1}\sum_{i,j}^{N_d,N_s} g_i^*\overline{u^j}\right\|_0 \\ &\le \frac{\max_{i,j}|\Gamma_{i,j}|}{\lambda}\sum_{i,j}^{N_d,N_s} \left\|g_i^* u^j\right\|_0 \end{aligned}$$

where we use the bound $\|\mathcal{K}^{-1}\|_{L^2(\Omega)\to L^2(\Omega)} \le 1/\lambda$, which can be obtained similar to the inequality (4.33). Then,

$$\|\alpha - \alpha^\lambda\|_0 \le \|\alpha\|_0 + \frac{\max_{i,j}|\Gamma_{i,j}|}{\lambda}\sum_{i,j}^{N_d,N_s} \left\|g_i^* u^j\right\|_0 .$$

Noting $u^j = u_0^j + u_s^j$, $\|g_i^* u_s^j\|_0 \le \|g_i^* u_0^j\|_0$, and recalling the bound for $\max_{i,j}|\Gamma_{i,j}|$ in (4.39), we get

$$\|\alpha - \alpha^\lambda\|_0 \le \|\alpha\|_0 + \frac{4}{\lambda}\max_{i,j}\|g_i^* u_0^j\|_0 \|\alpha\|_0 \sum_{i,j}^{N_d,N_s} \|g_i^* u_0^j\|_0.$$

Setting C_T as:

$$C_T = 1 + \frac{4}{\lambda} \max_{i,j} \|g_i^* u_0^j\|_0$$

leads to (4.44). □

Remarks:

1. The error e^λ originates from the error $\mathcal{I} - \mathcal{K}^{-1}\mathcal{A}$ in (4.46). (4.44) presents a loose but an explicit bound for the norm of the error e^λ in terms of $\|\alpha\|_\infty$. In practice, a sufficiently low value for λ will reduce the error e^λ when there is no noise.

2. Let γ^δ be the perturbed left-hand side γ of (5.19), such that $\|\gamma^\delta - \gamma\|_0 \leq \delta$. Then, an additional term is introduced to the bound in (4.44) due to this perturbation [57]:

$$\|\alpha - \alpha^\lambda\|_0 \leq \frac{\delta}{\lambda} + C_T\|\alpha\|_0. \tag{4.47}$$

4.4.3 The overall error

The following theorem provides a bound for the $L^1(\Omega)$ norm of the error in the reconstructed optical image $\tilde{\alpha}^\lambda$ with respect to the true absorptive perturbation α, resulting from Born approximation and Tikhonov regularization.

Theorem: A bound for the error in the reconstructed optical image $\tilde{\alpha}^\lambda$ with respect to the true heterogeneity α can be given by

$$\|\alpha - \tilde{\alpha}^\lambda\|_0 \leq C_T\|\alpha\|_0 + \frac{C_b}{\lambda} \sum_{i,j}^{N_d,N_s} \left(4\|g_i^* \tilde{\alpha}^\lambda\|_0 + \|\alpha\|_0 \|g_i^*\|_{L^\infty(\Omega)}\right) \|u_0^j \alpha\|_0 \tag{4.48}$$

Proof. The theorem is a result of (4.28), Lemma 3, and (4.44). □

4.5 Discussion

The presented error analysis, in particular Lemma 3, shows the circumstances under which the error in the reconstructed optical images due to Born approximation tends to increase. The orientation of the sources and detectors with respect to the support of the optical heterogeneity, the number of the sources and detectors, the size and magnitude of the heterogeneity are the factors that control the extent to which the linearization based on Born approximation is accurate. In this respect, Lemma 3 shows how the bound on the error varies with respect to these factors, rather than provide an exact quantitative measure of the resulting error in the reconstructed images. Provided there is *a priori* information about the support of the heterogeneity, the analysis can be used to place the sources and detectors such that the error in the reconstructed images resulting from linearization is reduced. One posterior approach can be to use only the specific sources-detectors in the image reconstruction, such that the error due to linearization by Born approximation is reduced. In addition, *a priori* information can be utilized to select an appropriate background absorption μ_{a0} such that the error in the reconstructed image due to Born approximation is minimized.

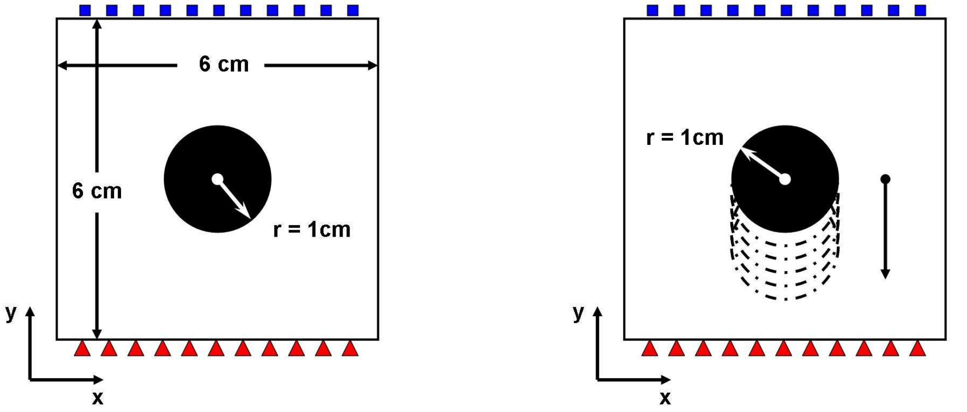
The analysis suggests some considerations for the nonlinear minimization methods based on iterative linearization [7] as well. These approaches convert the nonlinear inverse problem into a sequence of linear problems, where the solution of one linear problem is used to obtain the next. Therefore, the step length used to update the solution along a minimization direction is crucial in ensuring the validity of the linearization at each step. In other words, the new solution update following the linearization has to be within an interval such that the error introduced in the optical image stays in allowable limits. In order to avoid excessive error introduced by linearization, trust-region algorithms can be used to control the step length [60]. In this respect, Lemmas 2 and 3 can be utilized in the design of nonlinear optimization algorithms.

4.6 Numerical Experiments

We conduct a series of numerical experiments to demonstrate the implications of lemmas 2, 3, and the theorem. We perform our experiments in 2D for simplicity.

In the first simulation, we consider a series of image reconstructions to show the validity of lemma 2 and the theorem. In this context, we consider the imaging setup shown in figure 4.1(a) and change the absorption coefficient of the circular inclusion to observe the effect on the scattered field and imaging accuracy.

In the second simulation, we show how the proximity of the source to the optical heterogeneity affects the linearization accuracy. In this respect, we use the imaging setup shown in figure 4.1(b) and change the position of the circular heterogeneity in order to observe the effect on the scattered field and the accuracy of the reconstructed optical images.



(a) The optical domain and source-detector configuration for simulation study 1.

(b) The optical domain and source-detector configuration for simulation study 2. The radius of the circles is 0.75 cm.

Figure 4.1: The setups used for the simulation studies 1 and 2. The squares and triangles denote the detectors and sources, respectively.

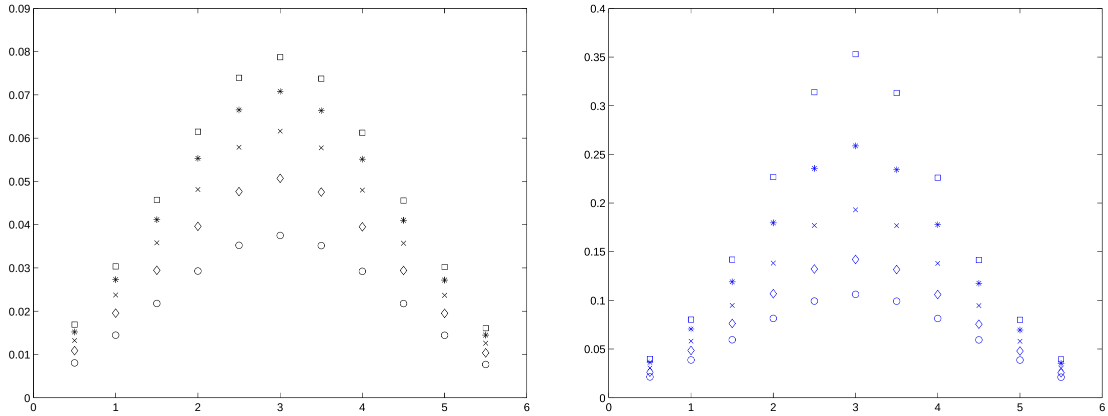
To discretize the boundary value problem in (4.5)-(4.6) and the associated adjoint problem (4.9)-(4.10), we use triangular finite elements with piecewise linear Lagrange basis functions [39]. For the discretization of the inverse problem (4.22), we use projection by collocation [39]. We apply Gaussian elimination method to solve the resulting finite-dimensional forward and inverse problems. We note that we make use of fine uniform meshes with 61 × 61 nodes to avoid any error in the reconstructed optical images resulting from discretization of the forward and inverse problems [40]. The regularization parameter is chosen as small as possible, yet large enough to enable robust image reconstructions. In this respect, an appropriate value for the regularization parameter is chosen based on experience and we set the regularization parameter λ to 10^{-7} in all experiments. Note that such a small λ value will reduce the error in the reconstructed image resulting from Tikhonov regularization.

4.6.1 Simulation 1

In this experiment, we consider the optical setup shown in figure 4.1(a). We use 11 sources and 11 detectors positioned on opposite sides. The source operating frequency is set to $\omega = 0$. The circular heterogeneity is embedded in a background with absorption coefficient $\mu_a = 0.04$ cm^{-1}. The diffusion coefficient of the medium $D(\mathbf{x}) = 0.0410$ is assumed to be constant for all $x \in \Omega \cup \partial\Omega$. The internal reflectance parameter is set to $a = 3$. In order to obtain a series of image reconstructions using the same setup, we change the absorption coefficient of the circular heterogeneity and consider five cases, where $\alpha = 0.04, 0.06, 0.08, 0.10$, and 0.12, respectively.

We first obtain the scattered field and compute the $L^2(\Omega)$ norm of the scattered field to verify lemma 2. Figure 4.2(a) reveals the dependence of the scattered field on the absorption coefficient of the optical heterogeneity. We see that the $L^2(\Omega)$ norm of the scattered field increases as the absorption coefficient of the heterogeneity increases. Furthermore, the $L^2(\Omega)$ norm of the scattered field is lower corresponding to the sources that are farther from the heterogeneity. This indicates the spatial dependence of the scattered field on the position of the heterogeneity with respect to the source as well. Figure 4.3(a) shows the cross-sections taken from $|u_0^j\alpha$ for $\alpha = 0.04$, $\alpha = 0.08$, and $\alpha = 0.12$, for which the jth source is located at $(3, 0)$. The cross-sections imply that $\|u_s^j\|_0$ gets larger as α increases, which is in consistence with the results shown in Figure 4.2(a).

In order to show the effect of Born approximation on the accuracy of the optical image reconstruction, we compare the reconstructed images with the true heterogeneity α. Figures 4.4(a)-4.4(c) and 4.4(e) show the reconstructed optical images. For brevity, we present the image reconstructions corresponding to the cases where $\alpha = 0.04, 0.08$, and 0.12cm^{-1}, respectively. We observe that the increase in the absorption coefficient of the circular heterogeneity leads to artifacts around the sources and detectors. Figures 4.4(b)-4.4(d) and 4.4(f) show the comparison of the



(a) The $L^2(\Omega)$ norm of the scattered optical field u_s corresponding to each of the 11 sources is shown for $\alpha = 0.04$ (circle), $\alpha = 0.06$ (diamond), $\alpha = 0.08$ (x-mark), $\alpha = 0.10$ (star), and $\alpha = 0.12$ (square).

(b) The $L^2(\Omega)$ norm of the scattered optical field u_s in simulation study 2, corresponding to each of the 11 sources. The results are shown for circular inclusions centered at $x = 3$ and $y = 2.75$ (circle), $y = 2.5$ (diamond), $y = 2.25$ (x-mark), $y = 2.0$ (star), and $y = 1.75$ (square).

Figure 4.2: The $L^2(\Omega)$ norm of the scattered field vs each source position (on x-axis) in each of the experiments in the first (a) and second (b) simulation studies.

cross-sectional cuts from the image reconstructions and the actual optical images. We see that the error between the reconstructed and the actual images increases as the absorption coefficient of the heterogeneity increases. More quantitatively, Table 3.2 shows the relationship between the error norm $\|\alpha - \tilde{\alpha}^\lambda\|$ and the absorption coefficient of the circular heterogeneity. $\|\alpha - \tilde{\alpha}^\lambda\|$ increases as the optical coefficient of the heterogeneity becomes larger. We note that the small value selected for the regularization parameter allows us to attribute the observed error to Born approximation. To see how the error $\|\alpha - \tilde{\alpha}^\lambda\|$ changes as α varies, in figure 4.3, we show the cross-sections of $|u_0^j\alpha|$ and $|g_i^*\tilde{\alpha}^\lambda|$ for the cases $\alpha = 0.04$, $\alpha = 0.08$, and $\alpha = 0.12$. As α increases, the increase in the value of $|u_0^j\alpha|$ and $|g_i^*\tilde{\alpha}^\lambda|$ is clear. Note that the bound (4.34) for the error resulting due to Born approximation is a function of $|u_0^j\alpha|$ and $|g_i^*\alpha^\lambda|$.

As a final remark, we note that the number of measurements and view angles can be increased to improve the quality of the reconstructed images. However, the objective of this simulation study is to show the effect of the absorption coefficient of the heterogeneity on the accuracy of the reconstructed images. Hence, using the same setup for each of the cases provides a fair comparison, which is sufficient to justify the proposed lemmas and the theorem.

4.6.2 Simulation 2

In this experiment, we consider the optical setup shown in figure 4.1(b). To simulate the optical data, we use the same source-detector configuration considered in the first simulation study. We simulate the optical data by solving the diffusion equation at $\omega = 0$ on a fine uniform grid with 61 nodes along x and y directions, where the refractive index mismatch parameter $a = 3$. The diffusion coefficient D is assumed to be constant and $D(\mathbf{x}) = 0.0410$ cm $x \in \Omega \cup \partial\Omega$. In order to assess the effect of the position of the heterogeneity on the accuracy of the image reconstructions, we gradually move the center of the circular inclusion towards the source side. We consider 5 cases, where the circular inclusion is centered at $(3, 2.75)$, $(3, 2.50)$, $(3, 2.25)$, $(3, 2)$ and $(3, 1.75)$, respectively.

We first obtain the scattered field and compute the $L^2(\Omega)$ norm of the scattered field to provide another verification of lemma 2. Figure 4.2(b) reveals the dependence of the scattered field on the proximity of the inclusion to the sources. We see that the $L^2(\Omega)$ norm of the scattered field increases as the center of the inclusion comes closer to the source side. Similar to the behavior observed in simulation study 1, for each case, we see that the $L^2(\Omega)$ norm of the scattered field is the highest for the source which is closest to the inclusion. We note that, as compared to its absorption coefficient, the position of the inclusion has a greater impact on the $L^2(\Omega)$ norm of the scattered field. In figure 4.5(a), we show the cross-sections taken from $|u_0^j\alpha|$ for three cases, in which the center of the circular inclusion is located at $(3, 2.75)$, $(3, 2.25)$, and $(3, 1.75)$, respectively. Noting that the jth source is positioned at $(3, 0)$, the increase in the value of $|u_0^j\alpha|$ as the inclusion gets closer to the source side is evident. These results are consistent with the results shown in figure 4.2(b).

Next, we compare the reconstructed images with the corresponding true heterogeneity α. We assume that the fine meshes used for the discretization of the forward and inverse problems and the low value of the regularization parameter ensure that the observed error is a result of Born approximation. Figures 4.6(a)-4.6(c) and 4.6(e) show the reconstructed optical images. For brevity, we present the image reconstructions corresponding to the cases where the center of the inclusion is located at $(3, 2.75), (3, 2.25)$, and $(3, 1.75)$, respectively. We observe that the artifacts around the sources tend to increase as the heterogeneity comes closer to the source side. Figures 4.6(b)-4.6(d) and 5.2(f) show the comparison of the cross-sectional cuts from the image reconstructions and the actual optical images. We see that the error between the reconstructed and the actual images increases as the center of the inclusion moves towards the source side. Table **??** provides a more quantitative result that displays the relationship between the error norm $\|\alpha - \tilde{\alpha}^\lambda\|$ and the position of the circular heterogeneity. $\|\alpha - \tilde{\alpha}^\lambda\|$ increases as the distance between the optical heterogeneity and the source side decreases. This is a direct validation of both lemmas 2 and 3, the latter of which shows that the error norm $\|\alpha - \tilde{\alpha}^\lambda\|$ depends on $\|u_0^j \alpha\|_0$. To see more clearly how the error $\|\alpha - \tilde{\alpha}^\lambda\|$ varies depending on the position of the circular heterogeneity, in figure 4.3, we show the cross-sections of $|u_0^j \alpha|$ and $|g_i^* \tilde{\alpha}^\lambda|$ for the cases in which the center of the circular inclusion is located at $(3, 2.75)$, $(3, 2.25)$, and $(3, 1.75)$. Noting that the jth source is positioned at $(3, 0)$, the increase in the value of $|u_0^j \alpha|$ is evident from figure 4.5(a) as the circular inclusion approaches the source side. On the other hand, there is slight decrease in the value of $g_i^* \tilde{\alpha}^\lambda$ as the inclusion moves away from the detector side. In this case, the increase in the error $\|\alpha - \tilde{\alpha}^\lambda\|$ is dominated by the increase in the term $\|u_0^j \alpha\|_0$ which multiplies $\|g_i^* \alpha^\lambda$ and $\|\alpha\|_\infty \|g_i^*\|_0$ in the error bound (4.34).

4.7 Conclusions

In this work, we presented an error analysis to show the effect of linearization of the inverse problem based on Born approximation on the accuracy of DOT image reconstructions. First, we derived an upper bound for the norm of the scattered optical field due to an absorptive perturbation. The bound was shown to depend on the spatial orientation of the optical heterogeneity and the optical field generated by the light sources. The bound provides a sensitive measure of the scattered field, since it takes the spatial variations of both the optical field and the optical heterogeneities into account. This makes the error bound significant especially in optical media with relatively small sizes. Next, we presented an error analysis to show the effect of zeroth-order Tikhonov regularization and Born approximation on the accuracy of DOT imaging. The analysis showed that the error in the reconstructed image due to Born approximation depends spatially on the optical heterogeneity, the optical field generated by the light sources, and the source-detector orientation. The error analysis indicates that there is a tradeoff between the norm of the scattered field and the accuracy of the inverse problem solution. Furthermore, the error analysis provides a good measure for the choice of step length in iterative linearization based nonlinear optimization methods, such as trust-region algorithms.

In our analysis, we assumed point sources that can be modelled by Gaussian functions with sufficiently low variance. We note that the analysis can be extended to arbitrary sources, provided they are $H^1(\Omega)$ bounded [39].

The error analysis presented in this work can be extended to show the effect of linearization on the accuracy of simultaneous reconstruction of scattering and absorption coefficients. Note that the presented error analysis is not limited to DOT, and can easily be adapted for similar inverse parameter estimation problems, in which Born approximation is applicable.

Table 4.1: The list of notation.

Notation:	**Explanation:**
Ω	Bounded domain in $\mathbb{R}^3$ with Lipschitz boundary
$\partial\Omega$	Lipschitz boundary of Ω
$\mathbf{x}$	Position vector in $\Omega \cup \partial\Omega$
$u_0^j(\mathbf{x})$	The incident field at $\mathbf{x}$ originating from background absorption μ_{a0}
$u_s^j(\mathbf{x})$	The scattered field at $\mathbf{x}$ due to a perturbation α over the background absorption μ_{a0}
$u^j(\mathbf{x})$	$(u_0^j + u_s^j)(\mathbf{x})$, the field at $\mathbf{x}$ due to the jth source
$g(\mathbf{x}, \acute{\mathbf{x}})$	The Green's function of (4.5)-(4.6) at $\mathbf{x}$ due to a point source located at $\acute{\mathbf{x}}$
$g_i^*(\mathbf{x})$	The solution of the adjoint problem (4.9)-(4.10) at $\mathbf{x}$ due to the adjoint source located at $\mathbf{x}_d^i$
$\Gamma_{i,j}$	Differential measurement at the ith detector due to the jth source
$\mathcal{A}_a$	The matrix-valued operator mapping $\alpha \in L^2(\Omega)$ to $\mathbb{C}^{N_d \times N_s}$
$\mathcal{A}_a^*$	The adjoint of $\mathcal{A}_a$ mapping from $\mathbb{C}^{N_d \times N_s}$ to $L^2(\Omega)$
$H_{i,j}(\mathbf{x})$	The kernel of $\mathcal{A}_a$ at $\mathbf{x}$
$H_{i,j}^*(\mathbf{x})$	The kernel of $\mathcal{A}_a^*$ at $\mathbf{x}$
$\gamma(\mathbf{x})$	$\mathcal{A}_a^*\Gamma$ at $\mathbf{x}$
$\tilde{\mathcal{A}}_a : L^2(\Omega) \to \mathbb{C}^{N_d \times N_s}$	The approximation to $\mathcal{A}_a$, resulting from Born approximation
$\tilde{\mathcal{A}}_a^* : \mathbb{C}^{N_d \times N_s} \to L^2(\Omega)$	The approximation to $\mathcal{A}_a^*$, resulting from Born approximation
$\tilde{H}_{i,j}(\mathbf{x})$, $\tilde{H}_{i,j}^*(\mathbf{x})$	The kernels of $\tilde{\mathcal{A}}_a$ and $\tilde{\mathcal{A}}_a^*$, respectively, at $\mathbf{x}$.
$\tilde{\gamma}(\mathbf{x})$	$\tilde{\mathcal{A}}_a^*\Gamma$ at $\mathbf{x}$
λ	The regularization parameter
$\alpha(\mathbf{x})$	Small perturbation over the background μ_a at $\mathbf{x}$
$\alpha^\lambda(\mathbf{x})$	Solution of the regularized inverse problem at $\mathbf{x}$ with exact operators
$\tilde{\alpha}^\lambda(\mathbf{x})$	Solution of the regularized inverse problem at $\mathbf{x}$ with operators resulting from Born approximation

Table 4.2: Definition of function spaces and norms.

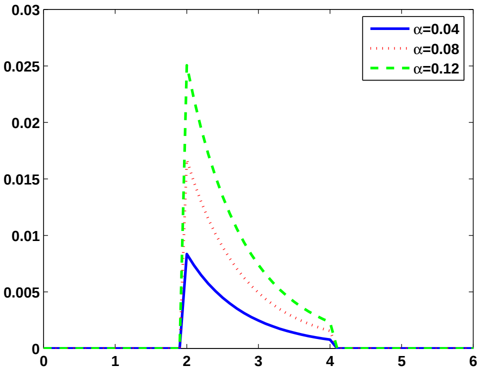
Notation	Explanation
$\overline{f}$	The complex conjugate of the function f
$C(\Omega)$	Space of continuous complex-valued functions on Ω
$C^k(\Omega)$	Space of complex-valued k-times continuously differentiable functions on Ω
$L^\infty(\Omega)$	$L^\infty(\Omega) = \{f \vert\ ess\sup_\Omega \vert f(\mathbf{x})\vert < \infty\ \}$
$L^p(\Omega)$	$L^p(\Omega) = \{f \vert\ (\int_\Omega \vert f(\mathbf{x})\vert^p d\mathbf{x})^{1/p} < \infty\ \}$, $p \in [1, \infty)$
$\Vert f\Vert_0$	The $L^2(\Omega)$ norm of f
$\Vert f\Vert_\infty$	The $L^\infty(\Omega)$ norm of f
$\Vert f\Vert_{L^p(\Omega)}$	The $L^p(\Omega)$ norm of f

Table 4.3: The error norm $\|\alpha - \tilde{\alpha}^\lambda\|_0$ measured for each image reconstruction in simulation study 1 (normalized by $\|\alpha - \tilde{\alpha}^\lambda\|_0$ for $\alpha = 0.040$ cm^{-1}.)

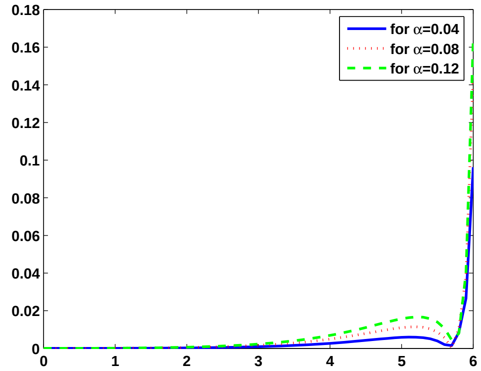
α:	0.040 cm^{-1}	0.060 cm^{-1}	0.080 cm^{-1}	0.10 cm^{-1}	0.120 cm^{-1}
$\Vert\alpha - \tilde{\alpha}^\lambda\Vert_0$:	1.0	1.3695	1.6895	1.9745	2.1908

Table 4.4: The error norm $\|\alpha - \tilde{\alpha}^\lambda\|_{L^1(\Omega)}$ measured for each calculation in simulation study 2 (normalized by $\|\alpha - \tilde{\alpha}^\lambda\|_0$ for $\alpha = 0.040$ cm^{-1} in the first experiment.

Circle center at:	(3,2.75)	(3,2.5)	(3,2.25)	(3,2.0)	(3,1.75)
$\Vert\alpha - \tilde{\alpha}^\lambda\Vert_0$:	2.2310	2.3656	2.4944	2.9119	3.5772



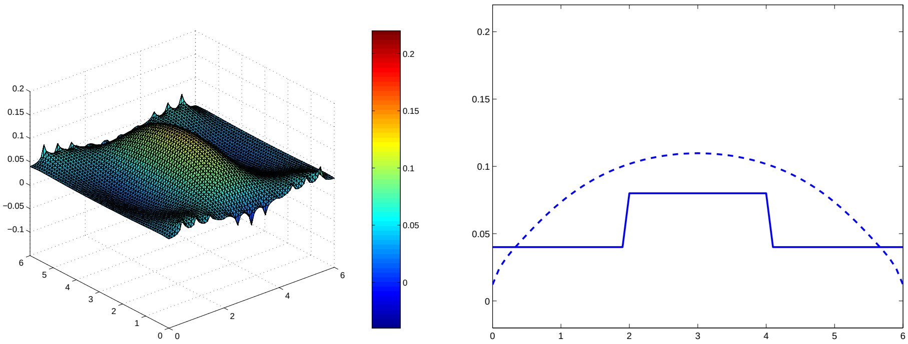
(a) The cross section of $|u_0^j\alpha|$ along $y-$axis at $x = 3$ in simulation study 1, for α values 0.04, 0.08, and 0.12.



(b) The cross section of $|g_i^*\tilde{\alpha}^\lambda|$ along $y-$axis at $x = 3$ in simulation study 1. The reconstructed functions $\tilde{\alpha}^\lambda$ correspond to the cases: $\alpha = 0.04$, $\alpha = 0.08$, and $\alpha = 0.12$.

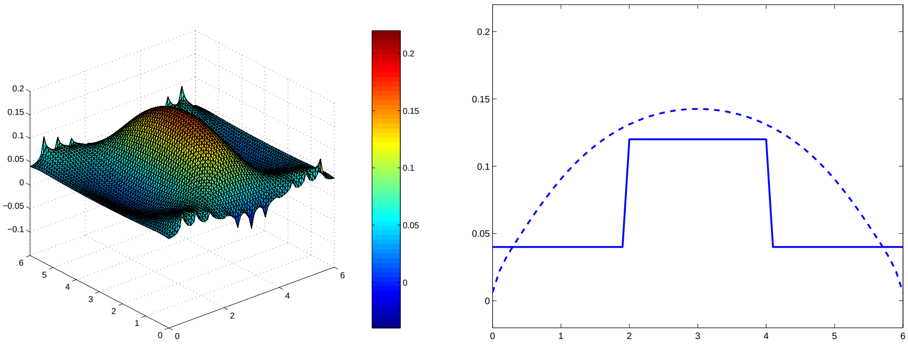
Figure 4.3: The cross-sections of $|u_0^j\alpha|$ and $|g_i^*\tilde{\alpha}^\lambda|$ in simulation study 1, which constitute the bound (4.34) for the error resulting from Born approximation. The jth source is located at $(3, 0)$ and the ith adjoint source is located at $(3, 6)$

.



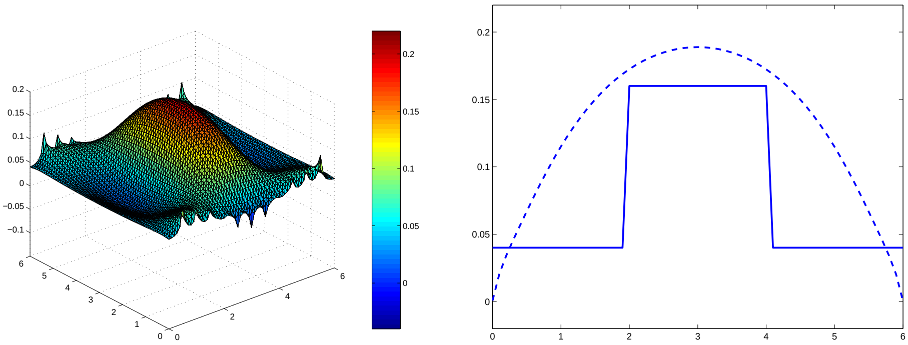
(a) The reconstructed absorption image for the case $\alpha = 0.04\text{cm}^{-1}$.

(b) The cross-sectional cuts taken from Figure 4.4(a) and the corresponding actual solution along y-axis at $x = 3$.



(c) The reconstructed absorption image for the case $\alpha = 0.08\text{cm}^{-1}$.

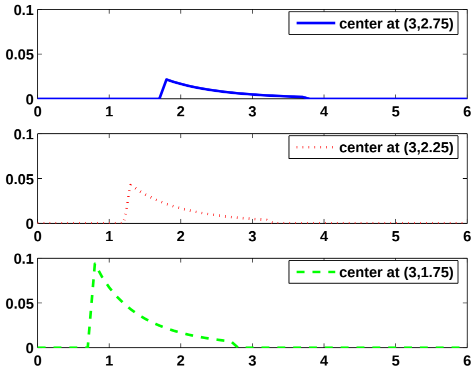
(d) The cross-sectional cuts taken from Figure 4.4(c) and the corresponding actual solution along y-axis at $x = 3$.



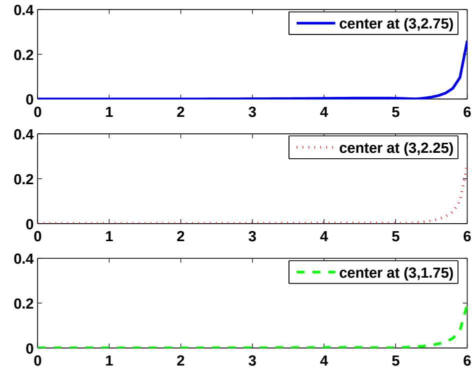
(e) The reconstructed absorption image for the case $\alpha = 0.12\text{cm}^{-1}$.

(f) The cross-sectional cuts taken from Figure 4.4(e) and the corresponding actual solution along y-axis at $x = 3$.

Figure 4.4: Samples of the reconstructed images in the first simulation study.

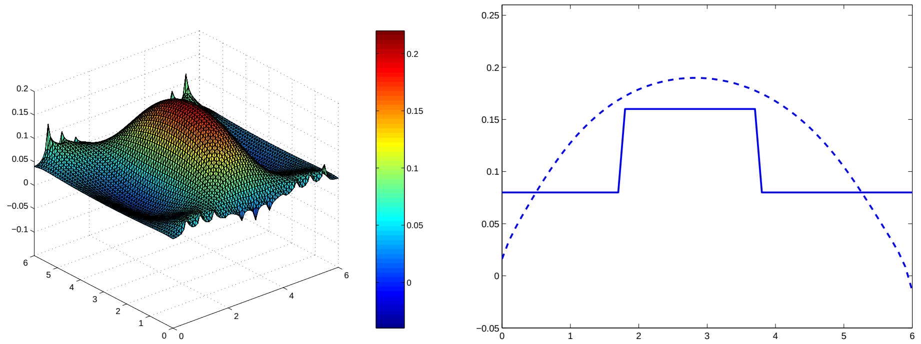


(a) The cross section of $|u_0^j \alpha|$ along $y-$axis at $x = 3$ in simulation study 2, for the circular inclusion centered at $(3, 2.75)$, $(3, 2.25)$, and $(3, 1.75)$.



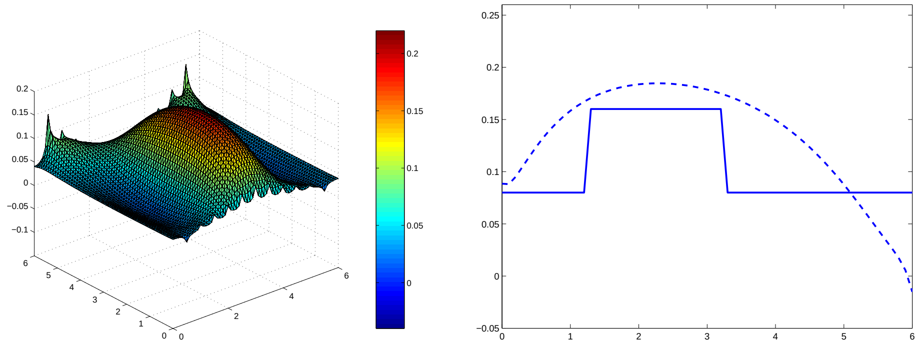
(b) The cross section of $|g_i^* \tilde{\alpha}^\lambda|$ along $y-$axis at $x = 3$ in simulation study 2. The reconstructed functions $\tilde{\alpha}^\lambda$ correspond to the cases in which the circular inclusion is centered at $(3, 2.75)$, $(3, 2.25)$, and $(3, 1.75)$.

Figure 4.5: The cross-sections of $|u_0^j \alpha|$ and $|g_i^* \tilde{\alpha}^\lambda|$ in simulation study 2, which constitute the bound (4.34) for the error resulting from Born approximation. The jth source is located at $(3, 0)$ and the ith adjoint source is located at $(3, 6)$.



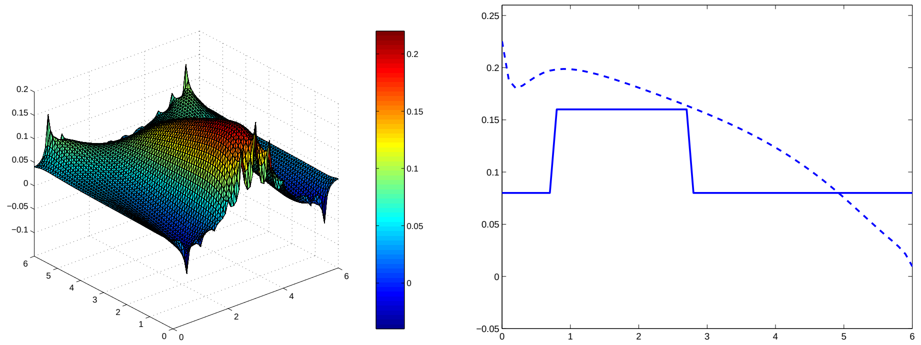
(a) Circular inclusion centered at $(x, y) = (3, 2.75)$.

(b) The cross-sectional cuts taken from Figure 4.6(a) and the corresponding actual solution along y-axis at $x = 3$.



(c) Circular inclusion centered at $(x, y) = (3, 2.25)$.

(d) The cross-sectional cuts taken from Figure 4.6(c) and the corresponding actual solution along y-axis at $x = 3$.



(e) Circular inclusion centered at $(x, y) = (3, 1.75)$.

(f) The cross-sectional cuts taken from Figure 4.6(e) and the corresponding actual solution along y-axis at $x = 3$.

Figure 4.6: Samples of the reconstructed images in the second simulation study.

CHAPTER 5

Effect of discretization on the accuracy of simultaneously reconstructed absorption and scattering images

Imaging in Diffuse Optical Tomography (DOT) is comprised of two interdependent stages which seek solutions to the forward and inverse problems. The forward problem is associated with describing the Near Infrared (NIR) light propagation, while the objective of the inverse problem is to estimate the unknown optical parameters from boundary measurements [4]. In this work, we model the forward problem by the diffusion equation in the frequency domain and the associated adjoint problem. For the inverse problem, we consider the simultaneous estimation of the optical diffusion and absorption coefficients.

A number of factors affect the accuracy of the DOT imaging: model accuracy (dependent on the light propagation model and/or linearization of the inverse problem), measurement noise, discretization, numerical algorithm efficiency, and inverse problem formulation. In this work, we focus on the effect of discretization on the accuracy of simultaneously reconstructed optical absorption and diffusion coefficients. In this respect, we extend our work in [39, 40]. First, we show the effect of forward problem discretization. Next, we show the effect of discretization of the inverse problem whose formulation uses the numerical solutions of the forward problem. Finally, we use the error analysis to devise novel adaptive mesh generation algorithms that reduce the error in the reconstructed optical images due to discretization for a given number of unknowns (i.e. for a given number of nodes in the adaptive meshes).

There has been extensive research on the estimation of discretization error in the solutions of partial differential equations (PDEs) [3, 10, 11, 13, 81, 83]. A somewhat different approach is followed in [71, 45] in which error in quantities of interest is related to the discretization of the second order elliptic partial differential equation. In the area of parameter estimation problems governed by PDEs, relatively little has been published. See for example [14] for an *a posteriori* error estimate for the Lagrangian in the inverse scattering problem for the time-dependent acoustic wave equation and [12] for a similar approach, and [62] for *a posteriori* error estimates for distributed elliptic optimal control problems.

In the area of DOT, it was numerically shown that the approximation errors resulting from the discretization of the forward problem can lead to significant errors in the reconstructed optical images [5]. However, an analysis regarding the error in the reconstructed optical images resulting from discretization has not been reported so far.

In this work, we model the forward problem by the frequency-domain diffusion equation. For the inverse problem, we focus on the simultaneous estimation of the absorption and diffusion coefficients. We consider the linear integral equation resulting from the iterative linearization of the inverse problem based on Born approximation and use zeroth order Tikhonov regularization to address the ill-posedness of the resulting integral equation. We use finite elements with first order Lagrange basis functions to discretize the forward and inverse problems and analyze the effect of the discretization on the reconstructed optical absorption and diffusion images. Our analysis shows that the error in the reconstructed optical images depends on the smoothness of the optical coefficients, the configuration of the source and detectors, the positions of the source and detectors with respect to locations of absorptive and diffusive heterogeneities, and on the regularization parameter in addition to the discretization error in the solution of each problem. In our analysis, we first consider the impact of the forward problem discretization when there is no discretization of the inverse problem, and provide a bound for the resulting error in the reconstructed optical images. Next, we analyze the effect of the discretization of the inverse problem whose formulation is based on the numerical (finite element) solutions of the forward problem and we obtain another bound for the resulting error in the reconstructed optical images. We see that each error bound comprises the discretization error in the corresponding problem solution, scaled spatially by the solutions of both problems. In addition, we notice that the error in the reconstruction of one optical parameter depends on how well the other optical parameter is discretized. Finally we note that the error analysis provides an insight into the so-called “inter-parameter crosstalk” [74, 44, 26] that originates from discretization (see figure 5.1 for an illustration). After summarizing the error analysis in two theorems, we use the error bounds presented in these theorems to devise novel adaptive mesh generation algorithms which are of $\mathcal{O}(N^{\Delta})$ complexity where N^{Δ} is the number of elements in the discretization mesh. In other words, the computational complexity of the proposed adaptive mesh generation algorithms does not exceed that of solving the resulting discrete forward and inverse problems.

The analysis presented in this work provides a means to identify and analyze the error in the simultaneously reconstructed optical images resulting from the linearization of the Lippmann-Schwinger type equations [25] using Born approximation, which will be an extension to our recent work [38]. Furthermore, the error analysis introduced in this paper is not limited to DOT, and can easily be extended for use in similar two-parameter inverse problems.

5.1 Forward and Inverse Problems

In this section, we describe the model for NIR light propagation and define the forward and inverse DOT problems. Table 5.1 provides a list of the notation and Table 5.2 provides the definition of function spaces and norms used throughout the paper. We note that we use calligraphic letters to denote the operators, e.g. $\mathcal{A}_a$, $\mathcal{I}$, $\mathcal{K}$ etc. For constant values we will use capital letters. The subscripts or superscripts a and b will be used respectively to denote the relevance to absorption and diffusion coefficients, which will be defined where they appear. The superscript $*$ denotes the adjoint and the rest of the superscripts and subscripts are clearly defined in the text. "tilde" will be used to denote a function or an operator is an approximation to its actual form. The spatial dependence of the functions will be implicitly assumed, without remarking it out. All relevant functions are defined with appropriate function spaces to avoid ambiguity. Vector quantities are in bold characters, such as $\mathbf{x}$.

5.1.1 Forward Problem

We use the following boundary value problem to model the NIR light propagation in a bounded domain $\Omega \subset \mathbb{R}^3$ with Lipschitz boundary $\partial\Omega$ [20, 4]:

$$-\nabla \cdot D(\mathbf{x})\nabla g_j(\mathbf{x}) + \left(\mu_a(\mathbf{x}) + \frac{i\omega}{c}\right) g_j(\mathbf{x}) = Q_j(\mathbf{x}) \quad \mathbf{x} \in \Omega, \tag{5.1}$$

$$g_j(\mathbf{x}) + 2aD(\mathbf{x})\frac{\partial g_j}{\partial n}(\mathbf{x}) = 0 \quad \mathbf{x} \in \partial\Omega, \tag{5.2}$$

where $g_j(\mathbf{x})$ is the photon density at $\mathbf{x} \in \Omega \cup \partial\Omega$ with frequency ω, Q_j is the j^{th} point source located at $\mathbf{x}_s^j$, $j = 1, \cdots, N_s$, where N_s is the number of sources. $D(\mathbf{x})$ is the diffusion coefficient and $\mu_a(\mathbf{x})$ is the absorption coefficient at $\mathbf{x}$, $i = \sqrt{-1}$, ω is the modulation frequency of the source, c is the speed of the light, $a = (1 + R)/(1 - R)$ where R is a parameter governing the internal reflection at the boundary $\partial\Omega$, and $\partial \cdot /\partial n$ denotes the directional derivative along the unit normal vector on the boundary. Note that we assume the diffusion coefficient is independent of the absorption coefficient and is isotropic. For the general anisotropic material, see [55].

The adjoint problem [4] associated with (5.1)-(5.2) is given by the following boundary value problem:

$$-\nabla \cdot D(\mathbf{x})\nabla g_i^*(\mathbf{x}) + \left(\mu_a(\mathbf{x}) - \frac{i\omega}{c}\right) g_i^*(\mathbf{x}) = 0 \quad \mathbf{x} \in \Omega, \tag{5.3}$$

$$g_i^*(\mathbf{x}) + 2aD(\mathbf{x})\frac{\partial g_i^*}{\partial n}(\mathbf{x}) = Q_i^*(\mathbf{x}) \quad \mathbf{x} \in \partial\Omega, \tag{5.4}$$

where Q_i^* is the adjoint source located at the i^{th} detector $\mathbf{x}_d^i$, $i = 1, \cdots, N_d$, where N_d is the number of detectors. We note that we approximate the point source Q_j in (5.1) and the adjoint source Q_i^* in (5.4) by Gaussian functions with sufficiently low variance, whose centers are located at $\mathbf{x}_s^j$ and $\mathbf{x}_d^i$, respectively. Note also that for any source $Q_j \in H^1(\Omega)$, our error analysis is valid.

In this work, we consider the finite-element approximations of the solutions of the forward problem. Hence, before we discretize the forward problem (see Section 5.2.1), we consider the variational formulations of (5.1)-(5.2) and (5.3)-(5.4) by multiplying (5.1) by a test function $\overline{\phi} \in H^1(\Omega)$ and integrating over Ω [20]:

$$\int_\Omega \left[\nabla\overline{\phi} \cdot D\nabla g_j + \overline{\phi}\left(\mu_a + \frac{i\omega}{c}\right) g_j - \overline{\phi} Q_j\right] d\mathbf{x} + \frac{1}{2a}\int_{\partial\Omega} \overline{\phi} g_j dl \quad = \quad 0, \tag{5.5}$$

where the boundary integral term results from the boundary condition (5.2). Let

$$\begin{aligned} A(\phi, g_j) &:= \int_\Omega \left[\nabla\overline{\phi} \cdot D\nabla g_j + \overline{\phi}\left(\mu_a + \frac{i\omega}{c}\right) g_j\right] d\mathbf{x}, \\ (\phi, Q_j) &:= \int_\Omega \overline{\phi} Q_j d\mathbf{x}, \\ \left\langle \phi, \frac{1}{2a} g_j \right\rangle &:= \frac{1}{2a}\int_{\partial\Omega} \overline{\phi} g_j dl. \end{aligned}$$

Then, we can express (D.6) equivalently by defining the sesquilinear form $b(\phi, g_j)$[39, 47]:

$$b(\phi, g_j) := A(\phi, g_j) + \left\langle \phi, \frac{1}{2a} g_j \right\rangle = (\phi, Q_j), \tag{5.6}$$

Similarly, the variational problem for (5.3)-(5.4) can be formulated by defining the sesquilinear form $b^*(\phi, g_i^*)$:

$$b^*(\phi, g_i^*) := A(\phi, g_i^*) + \left\langle \phi, \frac{1}{2a} g_i^* \right\rangle = \left\langle \phi, \frac{1}{2a} Q_i^* \right\rangle, \tag{5.7}$$

where in $A(\phi, g_i^*)$, ω is replaced by $-\omega$.

The sesquilinear forms $b(\phi, g_j)$, $b^*(\phi, g_i^*)$ are continuous and positive definite for bounded positive D and μ_a [47]. As a result, the variational problems (5.6) and (5.7) have unique solutions, which follows from the Lax-Milgram Lemma [20]. The solutions g_j and g_i^* of the variational problems (5.6) and (5.7) belong to $H^1(\Omega)$, which results from the H^1-boundedness of the Gaussian function that approximates the point source Q_j and the adjoint source Q_i^* [47].

Remark 1:

- In addition to above conditions, noting $Q_j, Q_i^* \in C(\overline{\Omega})$; the solutions g_j, g_i^* satisfy [27]

$$g_j, g_i^* \in W_1^\infty(\Omega). \tag{5.8}$$

- In this work, we address the forward problem by equations (5.6) and (5.7).

5.1.2 Inverse Problem

In this work, the objective of the inverse problem is to determine the unknown optical absorption and diffusion coefficients of a bounded optical domain. To address the nonlinear nature of the inverse DOT problem, we consider an iterative algorithm based on repetitive linearization of the inverse problem using first order Born approximation [4]. As a result, at each linearization step, the following linear integral equation relates the differential optical measurements to unknown small perturbations α and β on the absorption coefficient μ_a and the diffusion coefficient D, respectively, assuming $\beta = 0$ on $\mathbf{x} \in \partial\Omega$:

$$\begin{aligned}
\Gamma_{i,j} &= -\int_{\Omega} \left[\overline{g_i^*(\mathbf{x})} g_j(\mathbf{x})\alpha(\mathbf{x}) + \overline{\nabla g_i^*(\mathbf{x})} \cdot \nabla g_j(\mathbf{x})\beta(\mathbf{x})\right] d\mathbf{x} && (5.9)\\
&:= \int_{\Omega} \left[H^a_{i,j}(\mathbf{x})\alpha(\mathbf{x}) + H^b_{i,j}(\mathbf{x})\beta(\mathbf{x})\right] d\mathbf{x}\\
&:= \left(\begin{bmatrix} \mathcal{A}_a & \mathcal{B}_b \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix}\right)_{i,j}\\
&:= \left(\mathcal{K}_{ab}\sigma\right)_{i,j}, && (5.10)
\end{aligned}$$

where $\sigma = [\alpha\ \beta]^T \in L^2(\Omega) \times L^2(\Omega)$, $\mathcal{K}_{ab} = [\mathcal{A}_a\ \mathcal{B}_b] : L^2(\Omega) \times L^2(\Omega) \to \mathbb{C}^{N_d \times N_s}$, $H^a_{i,j}(\mathbf{x}) = -\overline{g_i^*(\mathbf{x})} g_j(\mathbf{x})$ is the (i,j)th kernel of the matrix valued operator $\mathcal{A}_a : L^2(\Omega) \to \mathbb{C}^{N_d \times N_s}$ at $\mathbf{x}$, and $H^b_{i,j}(\mathbf{x}) = -\overline{\nabla g_i^*(\mathbf{x})} \cdot \nabla g_j(\mathbf{x})$ is (i,j)th kernel of the matrix-valued operator $\mathcal{B}_b : L^2(\Omega) \to \mathbb{C}^{N_d \times N_s}$ at $\mathbf{x}$. g_j is the solution of (5.6) and $g_i^*(\mathbf{x})$ is the solution of (5.7), and $\Gamma_{i,j}$ is the (i,j)th entry in the vector $\Gamma \in \mathbb{C}^{N_d \times N_s}$, which represents the differential measurement at the i^{th} detector due to the j^{th} source [4, 39]. Thus,

$$\Gamma = K_{ab}\sigma. \tag{5.11}$$

Note that approximating Q_i^* in (5.4) by a Gaussian function centered at $\mathbf{x}_d^i$ implies that $\Gamma_{i,j}$ corresponds to the scattered optical field evaluated at $\mathbf{x}_d^i$, after filtering it by that Gaussian function [39]. Thus, the Gaussian approximation of the adjoint source models the finite size of the detectors. Similarly, approximating Q_j in (5.1) by a Gaussian function models the finite beam of the point source.

The operators $\mathcal{A}_a$ and $\mathcal{B}_b$ are bounded. In particular, for some $A_a, B_b > 0$ we have

$$\begin{aligned} \|\mathcal{A}_a\|_{L^2(\Omega)\to l^1} &\leq \sum_{i,j}^{N_d,N_s} \|g_i^* g_j\|_0 = A_a, && (5.12)\\ \|\mathcal{B}_b\|_{L^2(\Omega)\to l^1} &\leq \sum_{i,j}^{N_d,N_s} \|\nabla g_i^* \cdot \nabla g_j\|_0 = B_b, && (5.13) \end{aligned}$$

each of which is a consequence of remark 1.

Before we proceed with the boundedness of the integral operator $\mathcal{K}_{ab}$, let us define the norm $\|\sigma\|_{L^2(\Omega)\times L^2(\Omega)}$ as follows:

$$\|\sigma\|_{L^2(\Omega)\times L^2(\Omega)} := \left(\|\alpha\|_0^2 + \|\beta\|_0^2\right)^{1/2}.$$

Then, the operator $\mathcal{K}_{ab}$ is bounded by

$$\|\mathcal{K}_{ab}\|_{L^2(\Omega)\times L^2(\Omega)\to l^1} \leq A_a + B_b. \qquad (5.14)$$

Furthermore, the operators $\mathcal{A}_a$ and $\mathcal{B}_b$ are compact [39, 57]. Thus, for the given solution space $L^2(\Omega)$ for both α and β, (5.11) is ill-posed. To address the ill-posedness of (5.9), we regularize (5.11) with a zeroth order Tikhonov regularization.

5.1.3 Regularization of the inverse problem

Let $\mathcal{A}_a^* : \mathbb{C}^{N_d\times N_s} \to L^2(\Omega)$ and $\mathcal{B}_b^* : \mathbb{C}^{N_d\times N_s} \to L^2(\Omega)$ be the adjoint of the operators $\mathcal{A}_a$ and $\mathcal{B}_b$ defined respectively by

$$\begin{aligned} (\mathcal{A}_a^*\Theta)(\mathbf{x}) &:= \sum_{i,j}^{N_d,N_s} H_{i,j}^{a*}(\mathbf{x})\Theta_{i,j} := \sum_{i,j}^{N_d,N_s} -g_i^*(\mathbf{x})\overline{g_j(\mathbf{x})}\Theta_{i,j}, && (5.15)\\ (\mathcal{B}_a^*\Theta)(\mathbf{x}) &:= \sum_{i,j}^{N_d,N_s} H_{i,j}^{b*}(\mathbf{x})\Theta_{i,j} := \sum_{i,j}^{N_d,N_s} -\nabla g_i^*(\mathbf{x})\cdot\nabla\overline{g_j(\mathbf{x})}\Theta_{i,j}, && (5.16) \end{aligned}$$

for all $\Theta \in \mathbb{C}^{N_d \times N_s}$, where $H^{a*}_{i,j}$ and $H^{b*}_{i,j}$ are the (i,j)th kernels of $\mathcal{A}^*_a$ and $\mathcal{B}^*_b$, respectively. Then we define

$$\mathcal{K}^*_{ab} := \left[\begin{array}{cc} \mathcal{A}^*_a & \mathcal{B}^*_b \end{array} \right].$$

Let $\mathcal{A} := \mathcal{A}^*_a \mathcal{A}_a : L^2(\Omega) \to L^2(\Omega)$, $\mathcal{B} := \mathcal{B}^*_b \mathcal{B}_b : L^2(\Omega) \to L^2(\Omega)$, $\mathcal{A}_B := \mathcal{A}^*_a \mathcal{B} : L^2(\Omega) \to L^2(\Omega)$, and $\mathcal{B}_A := \mathcal{B}^*_b \mathcal{A}_a : L^2(\Omega) \to L^2(\Omega)$. Then,

$$\begin{aligned}
(\mathcal{A}\theta)(\mathbf{x}) &= \sum_{i,j}^{N_d,N_s} H^{a*}_{i,j}(\mathbf{x}) \int_\Omega H^a_{i,j}(\acute{\mathbf{x}})\theta(\acute{\mathbf{x}})d\acute{\mathbf{x}} := \int_\Omega \kappa_A(\mathbf{x},\acute{\mathbf{x}})\theta(\acute{\mathbf{x}})d\acute{\mathbf{x}}, \\
(\mathcal{B}\theta)(\mathbf{x}) &= \sum_{i,j}^{N_d,N_s} H^{b*}_{i,j}(\mathbf{x}) \int_\Omega H^b_{i,j}(\acute{\mathbf{x}})\theta(\acute{\mathbf{x}})d\acute{\mathbf{x}} := \int_\Omega \kappa_B(\mathbf{x},\acute{\mathbf{x}})\theta(\acute{\mathbf{x}})d\acute{\mathbf{x}}, \\
(\mathcal{A}_B\theta)(\mathbf{x}) &= \sum_{i,j}^{N_d,N_s} H^{a*}_{i,j}(\mathbf{x}) \int_\Omega H^b_{i,j}(\acute{\mathbf{x}})\theta(\acute{\mathbf{x}})d\acute{\mathbf{x}} := \int_\Omega \kappa_{AB}(\mathbf{x},\acute{\mathbf{x}})\theta(\acute{\mathbf{x}})d\acute{\mathbf{x}}, \\
(\mathcal{B}_A\theta)(\mathbf{x}) &= \sum_{i,j}^{N_d,N_s} H^{b*}_{i,j}(\mathbf{x}) \int_\Omega H^a_{i,j}(\acute{\mathbf{x}})\theta(\acute{\mathbf{x}})d\acute{\mathbf{x}} := \int_\Omega \kappa_{BA}(\mathbf{x},\acute{\mathbf{x}})\theta(\acute{\mathbf{x}})d\acute{\mathbf{x}},
\end{aligned}$$

for all $\theta \in L^2(\Omega)$ where

$$\kappa_A(\mathbf{x},\acute{\mathbf{x}}) := \sum_{i,j}^{N_d,N_s} H^{a*}_{i,j}(\mathbf{x})H^a_{i,j}(\acute{\mathbf{x}}), \quad \kappa_B(\mathbf{x},\acute{\mathbf{x}}) := \sum_{i,j}^{N_d,N_s} H^{b*}_{i,j}(\mathbf{x})H^b_{i,j}(\acute{\mathbf{x}}),$$
$$\kappa_{AB}(\mathbf{x},\acute{\mathbf{x}}) := \sum_{i,j}^{N_d,N_s} H^{a*}_{i,j}(\mathbf{x})H^b_{i,j}(\acute{\mathbf{x}}), \quad \kappa_{BA}(\mathbf{x},\acute{\mathbf{x}}) := \sum_{i,j}^{N_d,N_s} H^{b*}_{i,j}(\mathbf{x})H^a_{i,j}(\acute{\mathbf{x}}).$$

Note that $\kappa_{BA}(\mathbf{x},\acute{\mathbf{x}}) = \overline{\kappa_{AB}(\mathbf{x},\acute{\mathbf{x}})}$.

Let

$$\begin{aligned}
\mathcal{K} &:= \left[\begin{array}{cc} \mathcal{A} & \mathcal{A}_B \\ \mathcal{B}_A & \mathcal{B} \end{array} \right], && (5.17) \\
\gamma &:= \left[\begin{array}{c} \gamma_a \\ \gamma_b \end{array} \right] := \left[\begin{array}{c} \mathcal{A}^*_a\Gamma \\ \mathcal{B}^*_b\Gamma \end{array} \right], && (5.18)
\end{aligned}$$

where $\mathcal{K} : L^2(\Omega) \times L^2(\Omega) \to L^2(\Omega) \times L^2(\Omega)$. Then, using a zeroth order Tikhonov regularization, the following equation defines the inverse problem at each linearization step:

$$\begin{aligned} \gamma &= (\mathcal{K}+\mathcal{L})\,\sigma^\lambda \\ &:= \mathcal{T}\sigma^\lambda, \end{aligned} \tag{5.19}$$

where $\mathcal{T} := (\mathcal{K}+\mathcal{L}) : L^2(\Omega)\times L^2(\Omega) \to L^2(\Omega)\times L^2(\Omega)$, $\sigma^\lambda = [\alpha^\lambda\ \beta^\lambda]^T$ are approximations to α and β, respectively, and $\mathcal{L} : L^2(\Omega)\times L^2(\Omega) \to L^2(\Omega)\times L^2(\Omega)$ is given by

$$\mathcal{L} := \begin{bmatrix} \lambda_a \mathcal{I} & 0 \\ 0 & \lambda_b \mathcal{I} \end{bmatrix}, \tag{5.20}$$

where $\lambda_a, \lambda_b > 0$ and $\mathcal{I}$ is the identity operator. We finally note that a bound for $\mathcal{T}$ can be given by $\|\mathcal{T}\|_{L^2(\Omega)\times L^2(\Omega)\to L^2(\Omega)\times L^2(\Omega)} \le \|\mathcal{K}_{ab}\|^2_{L^2(\Omega)\times L^2(\Omega)\to l^1} + \max(\lambda_a, \lambda_b)$.

5.1.4 Existence and boundedness of the inverse operator

Consider the inverse problem formulation (5.19). Owing to the regularization term, the inverse operator $\mathcal{T}^{-1} : L^2(\Omega)\times L^2(\Omega) \to L^2(\Omega)\times L^2(\Omega)$ exists and by Lax-Milgram lemma [20] it is bounded by

$$\|\mathcal{T}^{-1}\|_{L^2(\Omega)\times L^2(\Omega)\to L^2(\Omega)\times L^2(\Omega)} \le \frac{1}{\min(\lambda_1, \lambda_2)}.$$

In particular, the operator $\mathcal{T}^{-1}$ can be viewed as a 2×2 matrix of operators $\mathcal{T}^{-1}_{ij} : L^2(\Omega) \to L^2(\Omega)$, $i, j = 1, 2$, i.e.:

$$\mathcal{T}^{-1} := \begin{pmatrix} \mathcal{T}^{-1}_{11} & \mathcal{T}^{-1}_{12} \\ \mathcal{T}^{-1}_{21} & \mathcal{T}^{-1}_{22} \end{pmatrix}. \tag{5.21}$$

We remark that the boundedness of $\mathcal{T}^{-1}$ is a result of the boundedness of the operators $\mathcal{T}^{-1}_{ij}$:

$$\|\mathcal{T}^{-1}_{ij}\|_{L^2(\Omega)\to L^2(\Omega)} \le \chi_{ij} \tag{5.22}$$

for some scalar value $\chi_{ij} > 0$, for $i, j = 1, 2$.

5.2 Discretization of the Forward and Inverse Problems

In this section, we first consider the variational formulations (5.6)-(5.7), and discuss the finite-element discretization of the forward problem. Next, we describe the discretization of the inverse problem (5.19).

5.2.1 Forward Problem Discretization

In this section, we consider the finite element discretization of (5.6) and (5.7), and use their solutions to approximate the kernels $\kappa_A, \kappa_B, \kappa_{AB}$, and κ_{BA}. As a result, we obtain finite dimensional approximations to $\mathcal{K}$ and γ in (5.17) and (5.18), respectively.

Let L_k be the first order Lagrange basis functions, and $Y_j \subset H^1(\Omega)$ be the finite-dimensional subspace spanned by L_k, $k = 1, \ldots, N_j$, where N_j is the dimension of the finite-dimensional subspace for the jth source, $j = 1, \ldots, N_s$. Similarly, we define $Y_i^* \subset H^1(\Omega)$ as the finite-dimensional subspace spanned by L_k, for $k = 1, \ldots, N_i^*$, where N_i^* is the dimension of the finite-dimensional subspace for the ith detector, $i = 1, \ldots, N_d$. In this representation, N_j and N_i^* denote that for each source and detector, the dimension of the finite-dimensional subspace can be different.

Replacing ϕ and g_j in (5.6) with their finite-dimensional counterparts $\Phi_j(\mathbf{x}) = \sum_{k=1}^{N_j} p_k L_k(\mathbf{x})$, $G_j(\mathbf{x}) = \sum_{k=1}^{N_j} c_k L_k(\mathbf{x})$; and replacing ϕ and g_i^* in (5.7) with $\Phi_i^*(\mathbf{x}) = \sum_{k=1}^{N_i^*} q_k L_k(\mathbf{x})$, $G_i^*(\mathbf{x}) = \sum_{k=1}^{N_i^*} d_k L_k(\mathbf{x})$ yields the matrix equations:

$$\mathbf{S}\mathbf{c}_j = \mathbf{q}_j, \tag{5.23}$$

$$\mathbf{S}^*\mathbf{d}_i = \mathbf{q}_i^*, \tag{5.24}$$

for $\mathbf{c}_j = [c_1, c_2, \cdots, c_{N_j}]^T$ and $\mathbf{d}_i = [d_1, d_2, \cdots, d_{N_i}]^T$. Here $\mathbf{S}$ and $\mathbf{S}^*$ are the finite element matrices and $\mathbf{q}_j$ and $\mathbf{q}_i^*$ are the load vectors resulting from the finite element discretization of (5.6) and (5.7).

The $H^1(\Omega)$ boundedness of the solutions g_j and g_i^* implies that the discretization error e_j and e_i^* in the finite element solutions G_j and G_i^* is bounded. Let $\{\Omega_{mj}\}$ denote the set of linear elements used to discretize (5.6) for $m = 1, \ldots, N_\Delta^j$; such that $\bigcup_m^{N_\Delta^j} \Omega_{mj} = \Omega$ for all $j = 1, \ldots N_s$. Similarly, let $\{\Omega_{ni}\}$ denote the set of linear elements used to discretize (5.7) for $n = 1, \ldots, N_\Delta^{*i}$; such that $\bigcup_n^{N_\Delta^{*i}} \Omega_{ni} = \Omega$ for all $i = 1, \ldots N_d$. Assuming the solutions g_j and g_i^* for all $j = 1, \cdots, N_s$ $i = 1, \cdots, N_d$ also satisfy $g_j, g_i^* \in H^2(\Omega)$, a bound for e_j and e_i^* on each finite element can be found by using the discretization error estimates ([20, 9]):

$$\|e_j\|_{0,mj} \leq C\|g_j\|_{2,mj} h_{mj}^2, \tag{5.25}$$

$$\|e_i^*\|_{0,ni} \leq C\|g_i^*\|_{2,ni} h_{ni}^2, \tag{5.26}$$

where C is a positive constant, $\|\cdot\|_{0,mj}$ ($\|\cdot\|_{0,ni}$) and $\|\cdot\|_{2,mj}$ ($\|\cdot\|_{2,ni}$) are respectively the L^2 and H^1 norms on Ω_{mj} (Ω_{ni}), and h_{mj} (h_{ni}) is the diameter of the smallest ball containing the finite element Ω_{mj} (Ω_{ni}) in the solution G_j (G_i^*). Similarly,

$$\|\nabla e_j\|_{0,mj} \leq C\|g_j\|_{2,mj} h_{mj}, \tag{5.27}$$

$$\|\nabla e_i^*\|_{0,ni} \leq C\|g_i^*\|_{2,ni} h_{ni}. \tag{5.28}$$

5.2.2 Approximation of $\mathcal{T}$ and γ with finite element solutions G_j and G_i^*

Following the discretization of the forward problem and the solution of the resulting discrete forward problem, we can rewrite the inverse problem formulation (5.19) by replacing g_j and g_i^* with G_j and G_i^* in $\mathcal{T}$ and γ. Consequently, we get the following inverse problem formulation, which is an approximation to the regularized inverse problem formulation in (5.19):

$$\tilde{\mathcal{T}} \tilde{\sigma}^\lambda = \tilde{\gamma}, \tag{5.29}$$

where

$$\tilde{\mathcal{T}} := \tilde{\mathcal{K}} + \mathcal{L}$$

$$:= \begin{bmatrix} \tilde{\mathcal{A}} & \tilde{\mathcal{A}}_B \\ \tilde{\mathcal{B}}_A & \tilde{\mathcal{B}} \end{bmatrix} + \begin{bmatrix} \lambda_a & 0 \\ 0 & \lambda_b \end{bmatrix} \tag{5.30}$$

$$\tilde{\gamma} := \begin{bmatrix} \tilde{\gamma}_a \\ \tilde{\gamma}_b \end{bmatrix} := \begin{bmatrix} \tilde{\mathcal{A}}_a^* \Gamma \\ \tilde{\mathcal{B}}_b^* \Gamma \end{bmatrix}. \tag{5.31}$$

are the approximations to $\mathcal{T}$ and γ, respectively, and $\tilde{\sigma}^\lambda$ is an approximation to the solution σ^λ as a result of the forward problem discretization:

$$\tilde{\sigma}^\lambda = \left[\tilde{\alpha}^\lambda \ \tilde{\beta}^\lambda\right]^T.$$

In (5.30), the kernels of the integral operators $\tilde{\mathcal{A}}$, $\tilde{\mathcal{B}}$, $\tilde{\mathcal{A}}_B$, and $\tilde{\mathcal{B}}_A$ are given respectively by

$$\tilde{\kappa}_A(\mathbf{x}, \acute{\mathbf{x}}) := \sum_{i,j}^{N_d, N_s} \tilde{H}_{i,j}^{a*}(\mathbf{x}) \tilde{H}_{i,j}^{a}(\acute{\mathbf{x}}), \quad \tilde{\kappa}_B(\mathbf{x}, \acute{\mathbf{x}}) := \sum_{i,j}^{N_d, N_s} \tilde{H}_{i,j}^{b*}(\mathbf{x}) \tilde{H}_{i,j}^{b}(\acute{\mathbf{x}}),$$
$$\tilde{\kappa}_{AB}(\mathbf{x}, \acute{\mathbf{x}}) := \sum_{i,j}^{N_d, N_s} \tilde{H}_{i,j}^{a*}(\mathbf{x}) \tilde{H}_{i,j}^{b}(\acute{\mathbf{x}}), \quad \tilde{\kappa}_{BA}(\mathbf{x}, \acute{\mathbf{x}}) := \sum_{i,j}^{N_d, N_s} \tilde{H}_{i,j}^{b*}(\mathbf{x}) \tilde{H}_{i,j}^{a}(\acute{\mathbf{x}}),$$

where

$$\begin{aligned} \tilde{H}_{i,j}^{a}(\mathbf{x}) &= -\overline{G_i^*(\mathbf{x})} G_j(\mathbf{x}), \\ \tilde{H}_{i,j}^{b}(\mathbf{x}) &= -\nabla \overline{G_i^*(\mathbf{x})} \cdot \nabla G_j(\mathbf{x}), \end{aligned}$$

Note that $\tilde{H}_{i,j}^{a*} = \overline{\tilde{H}_{i,j}^{a}}$ and $\tilde{H}_{i,j}^{b*} = \overline{\tilde{H}_{i,j}^{b}}$.

Similar to the operator $\mathcal{T}^{-1}$, $\tilde{\mathcal{T}}^{-1}$ can be interpreted as:

$$\tilde{\mathcal{T}}^{-1} := \begin{pmatrix} \tilde{\mathcal{T}}_{11}^{-1} & \tilde{\mathcal{T}}_{12}^{-1} \\ \tilde{\mathcal{T}}_{21}^{-1} & \tilde{\mathcal{T}}_{22}^{-1} \end{pmatrix} \tag{5.32}$$

where the operators $\tilde{\mathcal{T}}_{ij}^{-1}$ are bounded by:

$$\|\tilde{\mathcal{T}}_{ij}^{-1}\|_{L^2(\Omega) \to L^2(\Omega)} \leq \tilde{\chi}_{ij} \tag{5.33}$$

for some scalar value $\tilde{\chi}_{ij} > 0$, for $i, j = 1, 2$.

In the following section, we describe the discretization of the inverse problem (5.29) which uses the finite element approximations G_j and G_i^* of g_j and g_i^* in its formulation.

5.2.3 Discretization of the inverse problem

For the discretization of the inverse problem (5.29), we use projection by the Galerkin method. Below, we give the details of the Galerkin method.

Let $X^a, X^b \subset L^2(\Omega)$ denote the finite-dimensional subspaces spanned by first order Lagrange polynomials $\{L_1, \ldots, L_{N^a}\}$ and $\{L_1, \ldots, L_{N^b}\}$, associated with vertices located at x_p^a $p = 1, \cdots, N^a$ and x_r^b $r = 1, \cdots, N^b$, respectively, where N^a and N^b are the dimensions of X^a and X^b. Note that X^a and X^b are not necessarily identical.

Let $\{\Omega_t\}$, $t = 1, \cdots, N_\Delta^a$ denote a set of linear finite elements such that $\bigcup_t^{N_\Delta^a} \Omega_t = \Omega$ and $\{\Omega_u\}$ be a set of linear finite elements used for $u = 1, \cdots, N_\Delta^b$ such that $\bigcup_u^{N_\Delta^b} \Omega_u = \Omega$. Then, we express $\tilde{\sigma}_{n,m}^\lambda = [\tilde{\alpha}_n^\lambda \ \tilde{\beta}_m^\lambda]^T$ on these finite elements as

$$\tilde{\alpha}_n^\lambda(\mathbf{x}) = \sum_{k=1}^{N^a} a_k L_k(\mathbf{x}), \tag{5.34}$$

$$\tilde{\beta}_m^\lambda(\mathbf{x}) = \sum_{l=1}^{N^b} b_l L_l(\mathbf{x}). \tag{5.35}$$

Next consider the test function $\zeta = [\zeta_a \ \zeta_b]^T \in X^a \times X^b$ given by

$$\zeta_a(\mathbf{x}) = \sum_{k=1}^{N^a} c_k^a L_k(\mathbf{x}),$$

$$\zeta_b(\mathbf{x}) = \sum_{l=1}^{N^b} c_l^b L_l(\mathbf{x}).$$

Then, the Galerkin method approximates the solution of (5.29) by an element $\tilde{\sigma}_{n,m}^\lambda =$

$[\tilde{\alpha}_n^\lambda \ \tilde{\beta}_m^\lambda]^T \in X^a \times X^b$, which satisfies

$$\left(\tilde{\mathcal{T}}\sigma_{n,m}^\lambda, \zeta\right) = (\tilde{\gamma}, \zeta) \tag{5.36}$$

for all $\zeta \in (X^a \times X^b)$. We note that by Lax-Milgram theorem, a unique solution $\sigma_{n,m}^\lambda \in (X^a \times X^b)$ exists for (5.36) owing to the regularization which results in the positive-definiteness of the operator $\tilde{\mathcal{T}}$ [20, 57]. Equivalently, (5.36) can be interpreted as follows:

$$\mathcal{P}_{a,b}\tilde{\mathcal{T}}\sigma_{n,m}^\lambda = \mathcal{P}_{a,b}\tilde{\gamma}, \tag{5.37}$$

where $\mathcal{P}_{n,m}$ is the matrix of orthogonal projection operators

$$\mathcal{P}_{a,b} = \begin{pmatrix} \mathcal{P}_a & 0 \\ 0 & \mathcal{P}_b \end{pmatrix} \tag{5.38}$$

where $\mathcal{P}_a : L^2(\Omega) \to X^a$ and $\mathcal{P}_b : L^2(\Omega) \to X^b$ are the orthogonal projection operators [57]. We note that the following condition holds for $(\mathcal{P}_{a,b}\tilde{\mathcal{T}}) : X^a \times X^b \to X^a \times X^b$ (see proof of theorem 13.27 in [57]):

$$\|(\mathcal{P}_{a,b}\tilde{\mathcal{T}})^{-1}\mathcal{P}_{a,b}\|_{L^2(\Omega)\times L^2(\Omega)\to X^a\times X^b} \leq \frac{1}{\min(\lambda_a, \lambda_b)}. \tag{5.39}$$

5.2.4 Summary: The inverse problem and its approximations

In this work, we consider the regularized inverse problem in (5.19) as the baseline for the error analysis. In this respect, we first consider the effect of discretization of the forward problem on the optical imaging accuracy, thus consider the inverse problem (5.29), i.e.

$$\tilde{\mathcal{T}}\tilde{\sigma}^\lambda = \tilde{\gamma},$$

Next, to show the effect of inverse problem discretization, we project the above equation on the finite-dimensional subspaces $X^a \times X^b$, and consider the resulting inverse problem formulation (5.36)

$$\mathcal{P}_{a,b}\tilde{\mathcal{T}}\tilde{\sigma}^{\lambda}_{n,m} = \mathcal{P}_{a,b}\tilde{\gamma}.$$

Thus, we have three different inverse problem formulations:

1. The exact inverse problem formulation (5.19),
2. the inverse problem formulation (5.29) with the degenerate kernels, and
3. the full discrete inverse problem (5.36), which is the projection of (5.29) onto the finite dimensional subspaces $(X^a \times X^b) \subset (L^2(\Omega) \times L^2(\Omega))$ using Galerkin method.

5.3 Discretization-based Error Analysis

As a result of operator approximation and discretization of the inverse problem, the reconstructed images $\tilde{\sigma}^{\lambda}_n = [\tilde{\alpha}^{\lambda}_n \ \tilde{\beta}^{\lambda}_n]^T$ are approximations to the actual images $\sigma^{\lambda} = [\alpha^{\lambda} \ \beta^{\lambda}]$. Projecting the inverse problem onto finite-dimensional sub-spaces X^a and X^b and the discretization error in the solutions of the forward problem result in error in the reconstructed images. Therefore, the accuracy of the reconstructed image is challenged by the discretization schemes followed in the numerical solutions of the forward and inverse problems.

The error in the solution $\tilde{\sigma}^{\lambda}_{n,m}$ of (5.36) with respect to the actual solution σ^{λ} of (5.19) has two contributors: We write $\tilde{\sigma}^{\lambda}_{n,m} = \tilde{\sigma}^{\lambda} - e_{n,m}$, where $e_{n,m} = [e^a_n \ e^b_m]^T$ is the error resulting from projection of the inverse problem with the operator approximation and denote $\tilde{\sigma}^{\lambda} = \sigma^{\lambda} - \tilde{e}$, where $\tilde{e} = [\tilde{e}^a \ \tilde{e}^b]^T$ is the error due to forward problem discretization. As a result, we arrive at the following conclusion:

$$\tilde{\sigma}^{\lambda}_{n,m} = \tilde{\sigma}^{\lambda} - e_{n,m} = \sigma^{\lambda} - e_{n,m} - \tilde{e}. \tag{5.40}$$

Therefore, we can write an upper bound for the error $\sigma^{\lambda} - \tilde{\sigma}^{\lambda}_{n,m}$ as follows:

$$\|\sigma^{\lambda} - \tilde{\sigma}^{\lambda}_{n,m}\| = \|\sigma^{\lambda} - \tilde{\sigma}^{\lambda} + \tilde{\sigma}^{\lambda} - \tilde{\sigma}^{\lambda}_{n,m}\| = \|\tilde{e} + e_{n,m}\| \leq \|\tilde{e}\| + \|e_{n,m}\|. \tag{5.41}$$

5.3.1 Effect of forward problem discretization

The following theorem presents a bound for the $L^2(\Omega)$ norm of the error between the solution $\tilde{\sigma}^\lambda$ of (5.29) and the solution σ^λ of (5.19).

Theorem 1:

Let $\{\Omega_{mj}\}$ denote the set of linear elements used to discretize (5.6) for $m = 1, \ldots, N_\Delta^j$; such that $\bigcup_m^{N_\Delta^j} \Omega_{mj} = \Omega$, and h_{mj} be the diameter of the smallest ball that contains the element Ω_{mj} in the solution G_j, for all $j = 1, \ldots N_s$. Similarly, let $\{\Omega_{ni}\}$ denote the set of linear elements used to discretize (5.7) for $n = 1, \ldots, N_\Delta^{*i}$; such that $\bigcup_n^{N_\Delta^{*i}} \Omega_{ni} = \Omega$, and h_{ni} be the diameter of the smallest ball that contains the element Ω_{ni} in the solution G_i^*, for all $i = 1, \ldots N_d$. Assume further that the solutions g_j and g_i^* admit smoothness such that $g_j, g_i^* \in H^2(\Omega)$ and σ^λ is bounded, that is $\sigma^\lambda \in (L^\infty \times L^\infty)$. Let

$$\begin{aligned}
a(j,m) &:= \sum_{i=1}^{N_d} \|g_i^* \alpha^\lambda\|_{0,mj} \|g_j\|_{2,mj} \\
b(j,m) &:= \frac{\|\alpha\|_0 + \|\beta\|_0}{2} \sum_{i=1}^{N_d} \|g_i^*\|_{\infty,mj} \|g_j\|_{2,mj}, \\
c(j,m) &:= \sum_{i=1}^{N_d} \left\| |\nabla g_i^*| \beta^\lambda \right\|_{0,mj} \|g_j\|_{2,mj} \\
d(j,m) &:= \frac{\|\alpha\|_0 + \|\beta\|_0}{2} \sum_{i=1}^{N_d} \|\nabla g_i^*\|_{\infty,mj} \|g_j\|_{2,mj},
\end{aligned}$$

and

$$\begin{aligned}
a^*(i,n) &:= \sum_{j=1}^{N_s} \|g_j \alpha^\lambda\|_{0,ni} \|g_i^*\|_{2,ni} \\
b^*(i,n) &:= \frac{\|\alpha\|_0 + \|\beta\|_0}{2} \sum_{j=1}^{N_s} \|g_j\|_{\infty,ni} \|g_i^*\|_{2,ni}, \\
c^*(i,n) &:= \sum_{j=1}^{N_s} \left\| |\nabla g_j| \beta^\lambda \right\|_{0,ni} \|g_i^*\|_{2,ni} \\
d^*(i,n) &:= \frac{\|\alpha\|_0 + \|\beta\|_0}{2} \sum_{j=1}^{N_s} \|\nabla g_j\|_{\infty,ni} \|g_i^*\|_{2,ni}.
\end{aligned}$$

Given the *a priori* discretization error estimates (5.25)-(5.26) and a generic constant $C > 0$, a bound for the error between the solution α^λ and the solution $\tilde{\alpha}^\lambda$ of (5.19) due to the approximations $\tilde{\mathcal{T}}$ and $\tilde{\gamma}$ is given by:

$$\begin{aligned}
\|\alpha^\lambda - \tilde{\alpha}^\lambda\|_0 &\leq 2C \max_{i,j} \|g_i^* g_j\|_1 \\
\times &\left(\sum_{j=1}^{N_s} \sum_{m=1}^{N_\Delta^j} \left[(\tilde{\chi}_{11} + \tilde{\chi}_{12}) a(j,m) + \tilde{\chi}_{12} b(j,m)\right] h_{mj}^2 + \left[(\tilde{\chi}_{11} + \tilde{\chi}_{12}) c(j,m) + \tilde{\chi}_{12} d(j,m)\right] h_{mj} \right. \\
&\left. + \sum_{i=1}^{N_d} \sum_{n=1}^{N_\Delta^{*i}} \left[(\tilde{\chi}_{11} + \tilde{\chi}_{12}) a^*(i,n) + \tilde{\chi}_{12} b^*(i,n)\right] h_{ni}^2 + \left[(\tilde{\chi}_{11} + \tilde{\chi}_{12}) c^*(i,n) + \tilde{\chi}_{12} d^*(i,n)\right] h_{ni} \right)
\end{aligned}$$

and a bound for the error between the solution β^λ and the solution $\tilde{\beta}^\lambda$ of (5.19) due to the approximations $\tilde{\mathcal{T}}$ and $\tilde{\gamma}$ is given

$$\begin{aligned}
\|\beta^\lambda - \tilde{\beta}^\lambda\|_0 &\leq 2C \max_{i,j} \|g_i^* g_j\|_1 \\
\times &\left(\sum_{j=1}^{N_s} \sum_{m=1}^{N_\Delta^j} \left[(\tilde{\chi}_{21} + \tilde{\chi}_{22}) a(j,m) + \tilde{\chi}_{22} b(j,m)\right] h_{mj}^2 + \left[(\tilde{\chi}_{21} + \tilde{\chi}_{22}) c(j,m) + \tilde{\chi}_{22} d(j,m)\right] h_{mj} \right. \\
&\left. + \sum_{i=1}^{N_d} \sum_{n=1}^{N_\Delta^{*i}} \left[(\tilde{\chi}_{21} + \tilde{\chi}_{22}) a^*(i,n) + \tilde{\chi}_{22} b^*(i,n)\right] h_{ni}^2 + \left[(\tilde{\chi}_{21} + \tilde{\chi}_{22}) c^*(i,n) + \tilde{\chi}_{22} d^*(i,n)\right] h_{ni} \right)
\end{aligned}$$

***Proof*:** See proof in appendix (C.1) □.

Remark 2

1. Theorem 1 shows that the error in the reconstructed absorption image $\tilde{\alpha}^\lambda$ depends on the diffusive heterogeneity and the solutions of the forward problem. Similarly, the error in the reconstructed diffusion image $\tilde{\beta}^\lambda$ depends on the absorptive heterogeneity and the solutions of the forward problem. With these observations, theorem 1 suggests the use of meshes designed individually for the solutions G_j, $j = 1, \cdots, N_s$ and G_i^*, $i = 1, \cdots, N_d$. Note also that the position of the detectors with respect to the sources is another factor that affects the error bound in theorem 1.

2. Note that the conventional interpolation error estimates given in (5.25)-(5.26) and (5.27)-(5.28) depend on only the smoothness and support of g_j and g_i^* [20]. On the other hand, the error estimates in Theorem 1 show that the accuracy of the reconstructed images $\tilde{\alpha}\lambda$ and $\tilde{\beta}\lambda$ depend on the location of the absorptive and diffusive heterogeneities with respect to the sources and detectors, as well as on the bounds (5.25)-(5.26) and (5.27)-(5.28).

3. The parameters $\tilde{\chi}_{ij}$ $i, j = 1, 2$ affect the bounds on $\|\alpha^\lambda - \tilde{\alpha}^\lambda\|_0$ and $\|\beta^\lambda - \tilde{\beta}^\lambda\|_0$. Note that the parameters $\tilde{\chi}_{ij}$, $i, j = 1, 2$ depend on the regularization parameters λ_1, λ_2 and on the kernels of the operator $\mathcal{T}$. We also note that the kernels of $\mathcal{T}$ can be scaled to make $\tilde{\chi}_{ij}$ almost identical for all $i, j = 1, 2$ [**REF**]. Otherwise, the effect of forward problem discretization may be greater on one of the reconstructed optical coefficients as compared to the other one.

4. Increasing the number of sources and detectors increases the bounds on $\|\alpha^\lambda - \tilde{\alpha}^\lambda\|_0$ and $\|\beta^\lambda - \tilde{\beta}^\lambda\|_0$.

5.3.2 Effect of inverse problem discretization

In this section, we show the effect of inverse problem discretization on the optical imaging accuracy. In the analysis, we consider the inverse problem formulation and derive a bound for the $L^2(\Omega)$ norm of the error $e^{n,m}$ between the solution of (5.29) and the solution of (5.37).

Theorem 2:

Let $\{\Omega_t\}$ denote the set of linear elements used to discretize (5.34) for $t = 1, \ldots, N_\Delta^a$; such that $\bigcup_t^{N_\Delta^a} \Omega_t = \Omega$, and h_{ta} be the diameter of the smallest ball that contains the element Ω_t in the solution. Similarly, let $\{\Omega_u\}$ denote the set of linear elements used to discretize (5.35) for $u = 1, \ldots, N_\Delta^b$; such that $\bigcup_u^{N_\Delta^b} \Omega_u = \Omega$, and h_{ub} be the diameter of the smallest ball that contains the element Ω_u. Assume that the solutions $\tilde{\alpha}^\lambda$ and $\tilde{\beta}^\lambda$ are sufficiently smooth such that

$$\tilde{\alpha}^\lambda, \tilde{\beta}^\lambda \in H^1(\Omega).$$

Then,

$$\begin{aligned}\|\tilde{\alpha}^{\lambda}-\tilde{\alpha}_n^{\lambda}\|_0 \leq C(1+\lambda_a\pi_{11})\sum_{t=1}^{N_a^{\Delta}}\|\tilde{\alpha}^{\lambda}\|_{1,ta}h_{ta}+C\lambda_b\pi_{12}\sum_{u=1}^{N_b^{\Delta}}\|\tilde{\beta}^{\lambda}\|_{1,ub}h_{ub}\\ +C(\pi_{11}+\pi_{12})\max_{i,j}\|G_i^*G_j\|_1\\ \times\left(\sum_{t=1}^{N_a^{\Delta}}\sum_{i,j}^{N_d,N_s}\|G_i^*G_j\|_{0,ta}\|\tilde{\alpha}^{\lambda}\|_{1,ta}h_{ta}+\sum_{u=1}^{N_b^{\Delta}}\sum_{i,j}^{N_d,N_s}\|\nabla G_i^*\cdot\nabla G_j\|_{0,ub}\|\tilde{\beta}^{\lambda}\|_{1,ub}h_{ub}\right).\end{aligned}$$

and

$$\begin{aligned}\|\tilde{\beta}^{\lambda}-\tilde{\beta}_m^{\lambda}\|_0 \leq C(1+\lambda_b\pi_{22})\sum_{u=1}^{N_b^{\Delta}}\|\tilde{\beta}^{\lambda}\|_{1,ub}h_{ub}+C\lambda_a\pi_{21}\sum_{u=1}^{N_a^{\Delta}}\|\tilde{\alpha}^{\lambda}\|_{1,ta}h_{ta}\\ +C(\pi_{21}+\pi_{22})\max_{i,j}\|G_i^*G_j\|_1\\ \times\left(\sum_{t=1}^{N_a^{\Delta}}\sum_{i,j}^{N_d,N_s}\|G_i^*G_j\|_{0,ta}\|\tilde{\alpha}^{\lambda}\|_{1,ta}h_{ta}+\sum_{u=1}^{N_b^{\Delta}}\sum_{i,j}^{N_d,N_s}\|\nabla G_i^*\cdot\nabla G_j\|_{0,ub}\|\tilde{\beta}^{\lambda}\|_{1,ub}h_{ub}\right).\end{aligned}$$

***Proof*:** See appendix (C.2).

□

Remark 3

1. The theorem shows that the accuracy of the reconstructed image $\tilde{\alpha}_n^{\lambda}$ depends on the discretization scheme followed to discretize $\tilde{\beta}^{\lambda}$ as well as the discretization scheme followed to discretize $\tilde{\alpha}^{\lambda}$ itself.

2. Similarly, the theorem shows that the accuracy of the reconstructed image $\tilde{\beta}_m^{\lambda}$ depends on the discretization scheme followed to discretize $\tilde{\alpha}^{\lambda}$ as well as the discretization scheme followed to discretize $\tilde{\beta}^{\lambda}$ itself.

3. Theorem 2 shows the spatial dependence of the inverse problem discretization on the forward problem solution. The position of the detectors with respect to the sources is another factor that determines the extent of the error bound in theorem 2.

4. The parameters π_{ij} $i, j = 1, 2$ affect the bounds on $\|\tilde{\alpha}^\lambda - \tilde{\alpha}^\lambda_n\|_0$ and $\|\tilde{\beta}^\lambda - \tilde{\beta}^\lambda_m\|_0$. Note that, similar to $\tilde{\chi}_{ij}$, the parameters π_{ij}, $i, j = 1, 2$ depend on the regularization parameters λ_1, λ_2 and on the kernels of the operator $\tilde{\mathcal{T}}$. Similar to the kernels of $\mathcal{T}$, the kernels of $\tilde{\mathcal{T}}$ can be scaled to make π_{ij} almost equal **[REF]**.

5. Increasing the number of sources and detectors increases the bounds on $\|\alpha^\lambda - \tilde{\alpha}^\lambda_n\|_0$ and $\|\beta^\lambda - \tilde{\beta}^\lambda_m\|_0$.

5.3.3 Iterative linearization by Born approximation

In this section, we explore the error in the inverse problem solution within an iterative linearization approach.

The error analysis presented in this paper covers the error which results from the discretization of the forward and inverse problems. If the optical heterogeneities α and β are sufficiently low, then one iteration suffices to solve the inverse problem and the error analysis discussed in theorems 1 and 2 applies.

When iterative linearization is considered to address the nonlinearity of the inverse problem, we can make use of the error analysis at each linearized step as follows: Let $\sigma^\lambda_{(\tau)}$ and $\tilde{\sigma}^\lambda_{n,m(\tau)}$ be the actual solution of the regularized inverse problem (5.19) and the solution of (5.36) at the τth linearization step, respectively. At the end of the $(L-1)^{th}$ linearization step, the optical coefficient estimates at $\mathbf{x}$ are given by

$$\hat{\mu}_a^{(L-1)}(\mathbf{x}) = \mu_a^{(0)}(\mathbf{x}) + \sum_{\tau=1}^{L-1} \tilde{\alpha}^\lambda_{n(\tau)}(\mathbf{x}),$$
$$\hat{D}^{(L-1)}(\mathbf{x}) = D^{(0)}(\mathbf{x}) + \sum_{\tau=1}^{L-1} \tilde{\beta}^\lambda_{m(\tau)}(\mathbf{x}),$$

where $\tilde{\alpha}^\lambda_{n(\tau)}$ and $\tilde{\beta}^\lambda_{m(\tau)}$ possess errors due to discretization with respect to the actual solutions $\alpha^\lambda_{(\tau)}$, $\beta^\lambda_{(\tau)}$, and $\mu_a^{(0)}$ and D^0 are the initial guesses for the background absorption coefficient and diffusion coefficient, respectively.

In the next linearization, an error on the new solution updates $\hat{\mu}_a^{(L)}$, $\hat{D}^{(L)}$ will be introduced due to:

1. the error $(\tilde{\mathcal{T}} - \mathcal{T})^{(L-1)}$ in the operator $(\tilde{\mathcal{T}})^{(L-1)}$ and the error $(\tilde{\gamma} - \gamma)^{(L-1)}$ in $(\tilde{\gamma})^{(L-1)}$ resulting from the forward problem discretization,
2. the projection (inverse problem discretization).

Another error in the $(L-1)^{th}$ updates $\hat{\mu}_a^{(r-1)}$, $\hat{D}^{(L-1)}$ which is rather difficult to see arises in an indirect way: Note that $\hat{\mu}_a^{(L-1)}$ and $\hat{D}^{(L-1)}$ appear as the coefficients of the boundary value problems (5.1)-(5.2) and (5.3)-(5.4), which constitute the forward problem. An error in these coefficients implies perturbation in the solutions of (5.1)-(5.2) and (5.3)-(5.4). As a result, the finite element solutions G_j and G_i^* of the forward problem will possess error terms in addition to the discretization error, which leads to additional errors in the operator $\tilde{\mathcal{T}}$ and $\tilde{\gamma}$ other than the ones caused by the discretization of the forward problem.

As a result, the errors in $\hat{\mu}_a^{(L)}$ and $\hat{D}^{(L)}$ at the L^{th} iteration are bounded by:

$$\left\|\mu_a - \hat{\mu}_a^{(L)}\right\| = \left\| \sum_{\tau=1}^{L} \alpha_{(\tau)}^{\lambda} - \tilde{\alpha}_{n(\tau)}^{\lambda} \right\| \leq \sum_{\tau=1}^{L} \left\| \alpha_{(\tau)}^{\lambda} - \tilde{\alpha}_{n(\tau)}^{\lambda} \right\|,$$

$$\left\|D - \hat{D}^{(L)}\right\| = \left\| \sum_{\tau=1}^{L} \beta_{(\tau)}^{\lambda} - \tilde{\beta}_{m(\tau)}^{\lambda} \right\| \leq \sum_{\tau=1}^{L} \left\| \beta_{(\tau)}^{\lambda} - \tilde{\beta}_{m(\tau)}^{\lambda} \right\|,$$

assuming that the initial guesses $\mu_a^{(0)}$ and $D^{(0)}$ for the background absorption and diffusion coefficients are approximated accurately while solving the boundary value problems (5.1)-(5.2) and (5.3)-(5.4) at the first iteration, that is

$$\mu_a^{(0)}(\mathbf{x}) - \sum_{k=1}^{n} \mu_a^{(0)}(\mathbf{x}_k) L_k(\mathbf{x}) \to 0, \quad \mathbf{x} \in \Omega, \tag{5.42}$$

$$D^{(0)}(\mathbf{x}) - \sum_{k=1}^{m} D^{(0)}(\mathbf{x}_k) L_k(\mathbf{x}) \to 0, \quad \mathbf{x} \in \Omega. \tag{5.43}$$

Assuming that the error due to discretization which is induced indirectly is small as compared to the direct errors, theorems 1 and 2 can be considered to summarize the error due to discretization at each linearization step, which can be summed to obtain the overall error due to discretization at the end of iterative linearization.

5.4 Adaptive Mesh Generation

In this section, we discuss the design of adaptive mesh generation algorithms based on the error bounds presented in theorems 1 and 2. In this respect, we first consider the case in which the adaptive mesh generation aims to bound the error due to the discretization below some preset tolerance value and determine the requirements to establish such a tolerance value. This is followed by a practical realization of the adaptive mesh generation which is constrained by a predetermined maximum number of nodes in the final adaptive mesh.

5.4.1 Adaptive mesh generation for the forward problem

In this section, we make use of theorem 1 to devise an adaptive mesh generation algorithm for the discretization of the forward problem. The main objective of the discretization scheme is to uniformly distribute the error bound in theorem 1 onto the finite elements. We start with the following theorem that states the condition to ensure a pre-specified bound on the error resulting from forward problem discretization. Next, we propose a practical adaptive mesh generation algorithm that uses the results in theorem 1 and the theorem below.

Theorem 3:

Consider theorem 1. Let ε^f be the allowable bound on the error in $\tilde{\alpha}^\lambda$ and $\tilde{\beta}^\lambda$ with respect to respectively α^λ and β^λ, due to the discretization of the forward problem. Then, we define the parameter ε with the following equation:

$$\varepsilon := \frac{\varepsilon^f}{C \max_{i,j} \|g_i^* g_j\|_1} \left(\sum_{j=1}^{N_s} N_\Delta^j + \sum_{i=1}^{N_d} N_\Delta^{*i} \right)^{-1}, \tag{5.44}$$

where C is the generic constant in theorem 1. Let

$$H_{mj} := \frac{-\left[C_1 c(j,m) + C_2 d(j,m)\right] + \sqrt{\left[C_1 c(j,m) + C_2 d(j,m)\right]^2 + 4\left[C_1 a(j,m) + C_2 b(j,m)\right]\varepsilon}}{2[C_1 a(j,m) + C_2 b(j,m)]}, \tag{5.45}$$

and

$$H_{ni} := \frac{-\left[C_1 c^*(i,n) + C_2 d^*(i,n)\right] + \sqrt{\left[C_1 c^*(i,n) + C_2 d^*(i,n)\right]^2 + 4\left[C_1 a^*(i,n) + C_2 b^*(i,n)\right]\varepsilon}}{2[C_1 a^*(i,n), +C_2 b^*(i,n)]} \tag{5.46}$$

for some C_1, C_2, $0 < C_2 < C_1$ such that

$$(\tilde{\chi}_{11} + \tilde{\chi}_{12}), (\tilde{\chi}_{21} + \tilde{\chi}_{22}) \le C_1, \tag{5.47}$$

$$\tilde{\chi}_{12}, \tilde{\chi}_{22} \le C_2. \tag{5.48}$$

Then, if $h_{mj} > 0$ and $h_{ni} > 0$ satisfy

$$h_{mj} \le H_{mj}$$
$$h_{ni} \le H_{ni}$$

for m,j, $m = 1, \cdots, N_j^{\Delta}$, $j = 1, \cdots, N_s$ and for n,i, $n = 1, \cdots, N_i^{*\Delta}$, $i = 1, \cdots, N_d$, the following hold

$$\begin{aligned} \|\alpha^\lambda - \tilde{\alpha}^\lambda\|_0 &\le \varepsilon^f. \\ \|\beta^\lambda - \tilde{\beta}^\lambda\|_0 &\le \varepsilon^f. \end{aligned}$$

Proof. See proof in appendix (C.3).

□

In practice, H_{mj} and H_{ni} in (5.45)-(5.46) can not be computed since α, α^λ, β^λ, g_j, and g_i^* are unknown. However, H_{mj} and H_{ni} can be estimated by using approximations for the functions involved in these bounds, based on either *a priori* information or on the recent forward and inverse problem solution updates [40]. If there is no *a priori* information, α^λ and β^λ can be assumed to be spatially constant at the first linearization step. We note that for the initial mesh design, we use a model problem to compute the terms in the error bound relevant to the forward problem solution [40, 53, 85].

After each sweep of refinement, one has to update the ε value; on the other hand, after the first sweep of refinement, one can compute the bound H_{mj} and H_{ni} only for the new elements.

In case ε^f in (5.44) can not be chosen in prior, we consider a posterior approach. Let

$$\begin{aligned} \varepsilon^f(j,m) &:= [C_1 a(j,m) + C_2 b(j,m)]\, h_{mj}^2 + [C_1 c(j,m) + C_2 d(j,m)]\, h_{mj}, \quad (5.49) \\ \varepsilon^f(i,n) &:= [C_1 a^*(i,n) + C_2 b^*(i,n)]\, h_{ni}^2 + [C_1 c^*(i,n) + C_2 d^*(i,n)]\, h_{ni}. \quad (5.50) \end{aligned}$$

which are increasing functions in h_{mj} and h_{ni}, respectively. Then, for each source j (for each detector i) for the given initial h_{mj} (h_{ni}) values, compute $\varepsilon^f(j,m)$ ($\varepsilon^f(i,n)$) on each element and take the average to obtain $\overline{\varepsilon_j}$ ($\overline{\varepsilon_i}$). Then, for the elements for which $\varepsilon^f(j,m) > \overline{\varepsilon_j}$ ($\varepsilon^f(i,n) > \overline{\varepsilon_i}$), refine the element so that the new $\varepsilon^f(j,m)$ ($\varepsilon^f(i,n)$)for that element will be smaller. We note that, in this case, the algorithm has to be stopped when the number of nodes in the mesh (that is the number of unknowns in the discrete forward problem: N_j or N_i^*) exceeds the allowable number of nodes which is determined by the computational power at hand. To compute (5.49)-(5.50) one has to estimate the bounds $\tilde{\chi}_{ij}$, for $i,j = 1,2$ (see experiments section for an example) and g_j, g_i^* as well. g_j, g_i^* can be approximated by either the analytical solution on an unbounded domain [40] or their finite dimensional approximations (see algorithm 3). Finally we note that we use the approximation $\|g_j\|_{2,m} \approx \|g_j\|_{1,m}$ and $\|g_i^*\|_{2,n} \approx \|g_i^*\|_{1,n}$ for practical purposes, which is true for the finite element approximations G_j and G_i^*. In algorithm 3, we outline the practical adaptive mesh generation algorithm described above.

5.4.2 Adaptive mesh generation for the inverse problem

In this section, we use theorem 2 to devise an adaptive mesh generation algorithm for the discretization of the inverse problem formulation (5.29). Similar to the adaptive mesh generation algorithm for the numerical solution of the forward problem, the main objective of the discretization scheme is to uniformly distribute the error bound in theorem 2 onto the finite elements. We start with the following theorem that states the condition to ensure a pre-specified bound on the error

Algorithm 3 The pseudocode for the mesh generation algorithm for the forward problem.

```
◇ Generate an initial uniform mesh:
$(\Delta^j, N^j_\Delta)$, $\Delta = \bigcup_{m=1}^{N^j_\Delta}\{\Delta_m\}$ $\left((\Delta^{*i}, N^{*i}_\Delta), \Delta^{i*} = \bigcup_{n=1}^{N^{*i}_\Delta}\{\Delta_n\}\right)$
◇ Set the maximum number of nodes $N^f_{max}$
  while Number of nodes in $\Delta^j$ ($\Delta^{*i}$) less than $N^f_{max}$
    for each element $\Delta_m \in \Delta^j$ with mesh parameter $h_{mj}$ (for $\Delta_n \in \Delta^{*i}$ with $h_{ni}$)
      if first linearization
        ▪ Use analytical solutions for $g_j$ and $g_i^*$ and a priori anatomical
        information about $\alpha$ and $\beta$ to compute $\varepsilon^f(m,j)$ in (5.49) ($\varepsilon^f(n,i)$ in (5.50))
      else
        ▪ Use current solution updates $G_j$, $G_i^*$ and $\tilde{\alpha}_n^\lambda$, $\tilde{\beta}_m^\lambda$
        to compute $\varepsilon^f(m,j)$ in (5.49) ($\varepsilon^f(n,i)$ in (5.50))
    end
      ▪ Compute $\overline{\varepsilon}_j$ ($\overline{\varepsilon}_i$)
      ▪ Refine the elements with $\varepsilon^f(m,j) > \overline{\varepsilon_j}$ ($\varepsilon^f(n,i) > \overline{\varepsilon_i}$)
      ▪ Update the mesh $\Delta^j$ ($\Delta^{*i}$)
  end
◇ Solve for $G_j$ ($G_i^*$)
```

resulting from inverse problem discretization. Next, we propose a practical adaptive mesh generation algorithm that uses the results in theorem 2 and the theorem below.

Theorem 4:

Consider theorem 2. Let ξ^{inv} be the allowable bound on the error in $\tilde{\alpha}_n^\lambda$ and $\tilde{\beta}_m^\lambda$ with respect to respectively $\tilde{\alpha}^\lambda$ and $\tilde{\beta}^\lambda$, due to the inverse problem discretization. We define the parameter $\xi > 0$ with the following equation

$$\xi := \frac{\xi^{inv}}{C(N_a^\Delta + N_b^\Delta)}, \tag{5.51}$$

where $C > 0$ is the generic constant in theorem 2. Let

$$H_{ta} := \frac{\xi}{P_1\|\tilde{\alpha}^\lambda\|_{1,ta} + P_2 \max_{i,j}\|G_i^* G_j\|_1 \sum_{i,j}^{N_d,N_s}\|G_i^* G_j\|_{0,ta}\|\tilde{\alpha}^\lambda\|_{1,ta}}, \tag{5.52}$$

and

$$H_{ub} := \frac{\xi}{P_1 \|\tilde{\beta}^\lambda\|_{1,ub} + P_2 \max_{i,j} \|G_i^* G_j\|_1 \sum_{i,j}^{N_d,N_s} \|G_i^* G_j\|_{0,ub} \|\tilde{\beta}^\lambda\|_{1,ub}} \tag{5.53}$$

for some $P_1, P_2 > 0$ such that

$$(1 + \lambda_a \pi_{11}), (1 + \lambda_b \pi_{22}) \leq P_1 \tag{5.54}$$

$$(\pi_{11} + \pi_{12}), (\pi_{21} + \pi_{22}) \leq P_2 \tag{5.55}$$

Then, if if $h_{ta} > 0$ and $h_{ub} > 0$ satisfy

$$\begin{aligned} h_{ta} &\leq H_{ta}, \\ h_{ub} &\leq H_{ub}, \end{aligned}$$

for all t,u, $t = 1, \cdots, N_a^\Delta$, $u = 1, \cdots, N_b^\Delta$, then the error in the reconstructed images $\tilde{\alpha}_n^\lambda$, $\tilde{\beta}_m^\lambda$ due to inverse problem discretization is bounded by

$$\begin{aligned} \|\tilde{\alpha}^\lambda - \tilde{\alpha}_n^\lambda\|_0 &\leq \xi^{inv}, \\ \|\tilde{\beta}^\lambda - \tilde{\beta}_m^\lambda\|_0 &\leq \xi^{inv}. \end{aligned}$$

Proof. Substituting the bounds (5.52)-(5.53) in theorem 2 leads to the theorem.

□

In practice, H_{ta}, H_{ub} in (5.52) and (5.53) can not be computed since $\tilde{\alpha}^\lambda$, $\tilde{\beta}^\lambda$. Similar to the approach described in Section 3.3.1, H_{ta} and H_{ub} can be estimated by using approximations for the functions involved in these bounds, based on either *a priori* information or on the recent inverse problem solution updates. If there is no *a priori* information, $\tilde{\alpha}^\lambda$ and $\tilde{\beta}^\lambda$ can be assumed to be spatially constant at the first linearization step.

After each sweep of refinement, one has to update the ξ value; on the other hand, after the first sweep of refinement, one can compute the bound H_{ta} and H_{ub} only for the new elements.

In case ξ^{inv} in (5.51) can not be chosen in prior, we consider a posterior approach. Let

$$\xi_a^{inv}(t) := \left(P_1 \|\tilde{\alpha}^\lambda\|_{1,ta} + P_2 \max_{i,j} \|G_i^* G_j\|_1 \sum_{i,j}^{N_d,N_s} \|G_i^* G_j\|_{0,ta} \|\tilde{\alpha}^\lambda\|_{1,ta} \right) h_{ta} \quad (5.56)$$

$$\xi_b^{inv}(u) := \left(P_1 \|\tilde{\beta}^\lambda\|_{1,ub} + P_2 \max_{i,j} \|G_i^* G_j\|_1 \sum_{i,j}^{N_d,N_s} \|G_i^* G_j\|_{0,ub} \|\tilde{\beta}^\lambda\|_{1,ub} \right) h_{ub}. \quad (5.57)$$

Then, for the discretization of the absorptive heterogeneity (of the diffusive heterogeneity) for the given initial h_{ta} (h_{ub}) values, compute $\xi_a^{inv}(t)$ ($\xi_b^{inv}(u)$) on each element and take the average to obtain $\overline{\xi_a^{inv}}$ ($\overline{\xi_b^{inv}}$). Then, for the elements for which $\xi_a^{inv}(t) > \overline{\xi_a^{inv}}$ ($\xi_b^{inv}(u) > \overline{\xi_b^{inv}}$), refine the element so that the new $\xi_a^{inv}(t)$ ($\xi_b^{inv}(u)$) for that element will be smaller. We note that, in this case, the algorithm has to be stopped when the number of nodes in the mesh exceeds the maximum number of nodes reserved for the inverse problem discretization. To compute (5.58)-(5.59), one has to estimate the bounds Π_{ij}, for $i, j = 1, 2$ as well (see experiments section for an example). In algorithm 4, we outline the practical adaptive mesh generation algorithm for the inverse problem discretization described above.

5.4.3 Computational complexity of the adaptive mesh generation algorithms

In this section, we briefly discuss the computational complexity of the adaptive mesh generation algorithms described in the previous sections. We first start with the adaptive mesh generation algorithm for the forward problem discretization, which is followed by the adaptive mesh generation algorithm for the inverse problem discretization.

Using triangular finite elements with first order Lagrange basis functions and an analytical (exact) integration on each finite element, the number of multiplications required to compute the L^2 or H^1 norm of a finite dimensional function (or two functions) on each triangular element is of $\mathcal{O}(1)$ complexity [40]. Then, the computational complexity of computing the bound (5.49) becomes $\mathcal{O}(N_d N_s)$ for each element. For all elements, the computational complexity amounts to $\mathcal{O}(N_\Delta^j N_d N_s)$. Similarly,

Algorithm 4 The pseudocode for the mesh generation algorithm at every linearization step of the inverse problem.

```
◇ Generate an initial uniform mesh
$(\Delta^a, N^a_\Delta)$, $\Delta^a = \bigcup_{t=1}^{N^a_\Delta} \{\Delta_t\}$ $\left((\Delta^b, N^b_\Delta), \Delta^b = \bigcup_{u=1}^{N^b_\Delta} \{\Delta_u\}\right)$
◇ Set the maximum number of nodes $N^a_{max}$ $(N^b_{max})$
  while Number of nodes $N^a$ $(N^b)$ less than $N^a_{max}$ $(N^b_{max})$
    for each element $\Delta_t \in \Delta^a$ with mesh parameter $h_{ta}$ (for $\Delta_u \in \Delta^b$ with $h_{ub}$)
      if first linearization
        . Use current solution updates $G_j$ and $G_i^*$ and a priori information
        about $\alpha$ and $\beta$ to compute $H_{ta}$ in (5.52) ($H_{ub}$ in (5.53))
      else
        . Use current solution updates $G_j$ and $G_i^*$ and $\tilde{\alpha}_n^\lambda$ $(\tilde{\beta}_m^\lambda)$
        to compute $\xi_a^{inv}(t)$ in (5.58) ($\xi_b^{inv}(t)$ in (5.59))
    end
      . Compute $\overline{\xi_a^{inv}}$ $\left(\overline{\xi_b^{inv}}\right)$
      . Refine the elements with $\xi_a^{inv}(t) > \overline{\xi_a^{inv}}$ $\left(\xi_b^{inv}(u) > \overline{\xi_b^{inv}}\right)$
      . Update the mesh $\Delta^a$ $(\Delta^b)$
  end
◇ Solve for $\tilde{\alpha}_n^\lambda$ and $\tilde{\beta}_m^\lambda$.
```

one can obtain the same computational complexity to compute the bound (5.50).

It is possible to even reduce the complexity by making the following modifications in $a(j,m), b(j,m), c(j,m), d(j,m)$ and in $a^*(j,m), b^*(j,m), c^*(j,m), d^*(j,m)$ in theorem 1:

$$
\begin{aligned}
a(j,m) &\approx \left\| \sum_{i=1}^{N_d} g_i^* \alpha^\lambda \right\|_{0,mj} \|g_j\|_{2,mj} \\
b(j,m) &\approx \frac{\|\alpha\|_0 + \|\beta\|_0}{2} \left\| \sum_{i=1}^{N_d} g_i^* \right\|_{\infty,mj} \|g_j\|_{2,mj}, \\
c(j,m) &\approx \left\| \sum_{i=1}^{N_d} |\nabla g_i^*| \beta^\lambda \right\|_{0,mj} \|g_j\|_{2,mj} \\
d(j,m) &\approx \frac{\|\alpha\|_0 + \|\beta\|_0}{2} \left\| \sum_{j=1}^{N_s} \nabla g_i^* \right\|_{\infty,mj} \|g_j\|_{2,mj},
\end{aligned}
$$

and

$$\begin{aligned}
a^*(i,n) &\approx \left\| \sum_{j=1}^{N_s} g_j \alpha^\lambda \right\|_{0,ni} \|g_i^*\|_{2,ni} \\
b^*(i,n) &\approx \frac{\|\alpha\|_0 + \|\beta\|_0}{2} \left\| \sum_{j=1}^{N_s} g_j \right\|_{\infty,ni} \|g_i^*\|_{2,ni}, \\
c^*(i,n) &\approx \left\| \sum_{j=1}^{N_s} |\nabla g_j| \beta^\lambda \right\|_{0,ni} \|g_i^*\|_{2,ni} \\
d^*(i,n) &\approx \frac{\|\alpha\|_0 + \|\beta\|_0}{2} \left\| \sum_{j=1}^{N_s} \nabla g_j \right\|_{\infty,ni} \|g_i^*\|_{2,ni}.
\end{aligned}$$

Then, the computational complexity of adaptive mesh generation for each source or detector becomes of $\mathcal{O}(N_s N_\Delta) + \mathcal{O}(N_d N_\Delta)$ complexity.

Next, we discuss the computational complexity of adaptive mesh generation for the discretization of the inverse problem. In a similar way as above, the computational complexity of computing (5.58) and (5.59) can be shown to be of $\mathcal{O}(N_d N_s N_\Delta)$. Further reduction is possible by making the following approximations in (5.58) and (5.59):

$$\xi_a^{inv}(t) \approx \left(P_1 \|\tilde{\alpha}^\lambda\|_{1,ta} + P_2 \max_{i,j} \|G_i^* G_j\|_1 \left\| \sum_{i,j}^{N_d,N_s} G_i^* G_j \right\|_{0,ta} \|\tilde{\alpha}^\lambda\|_{1,ta} \right) h_{ta} \quad (5.58)$$

$$\xi_b^{inv}(u) \approx \left(P_1 \|\tilde{\beta}^\lambda\|_{1,ub} + P_2 \max_{i,j} \|G_i^* G_j\|_1 \left\| \sum_{i,j}^{N_d,N_s} G_i^* G_j \right\|_{0,ub} \|\tilde{\beta}^\lambda\|_{1,ub} \right) h_{ub}. \quad (5.59)$$

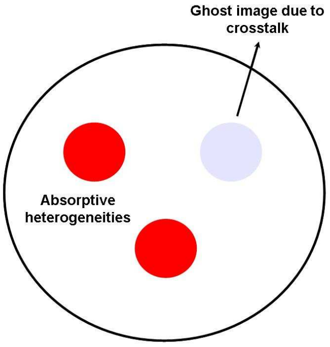
Then, the computational cost of adaptive mesh generation for the discretization of the inverse problem becomes of $\mathcal{O}(N_\Delta^a) + \mathcal{O}(N_\Delta^b)$ complexity.

5.5 Conclusion

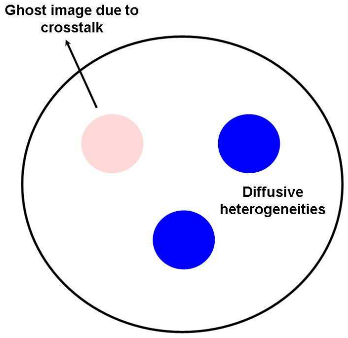
In this work, we presented an error analysis to show the relationship between the error in the simultaneously reconstructed optical absorption and diffusion coefficient images and the discretization of the forward and inverse problems.

We summarized the results of the error analysis in two theorems which provide an insight into the effect of forward and inverse problem discretizations on the accuracy of diffuse optical imaging. These theorems show that the error in the reconstructed optical images due to the discretization of each problem is bounded by roughly the multiplication of the discretization error in the corresponding solution and the solution of the other problem. One important implication of the error bounds is the dependence of the error in the reconstruction of one optical parameter (say the absorption coefficient) on the discretization of the other optical parameter (say the diffusion coefficient).

Based on the error analysis, we developed two new adaptive mesh generation algorithms, one for the forward and one for the inverse problem, which take into account the interdependence between the solutions of the two problems.



(a) Reconstructed absorption image with crosstalk. Red circles indicate the reconstructed absorptive heterogeneities and the light blue circle on the top right is the ghost image introduced as a result of crosstalk from the diffusion image.



(b) Reconstructed diffusion image with crosstalk. Blue circles indicate the reconstructed absorptive heterogeneities and the pink circle on the top left is the ghost image introduced as a result of crosstalk from the absorption image.

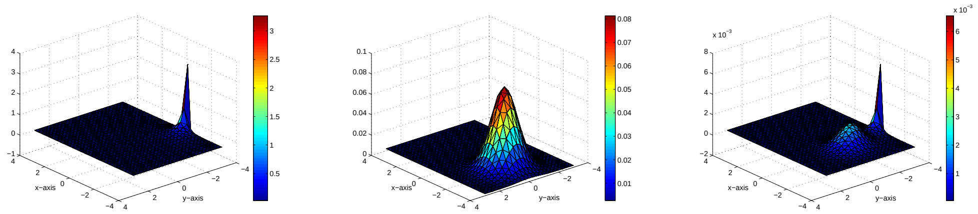
Figure 5.1: An example of crosstalk in simultaneous reconstruction of absorption and diffusion coefficients.

Table 5.1: Definition of variables, functions, and operators.

Notation	Explanation
Ω	Bounded domain in $\mathbb{R}^3$ with Lipschitz boundary
$\partial\Omega$	Lipschitz boundary of Ω
$\mathbf{x}$	Position vector in $\Omega \cup \partial\Omega$
$g_j(\mathbf{x})$	Solution of the diffusion equation at $\mathbf{x}$ for the j^{th} point source located at $\mathbf{x}_s^j$
$g_i^*(\mathbf{x})$	Solution of the adjoint problem at $\mathbf{x}$ for the i^{th} adjoint source located at $\mathbf{x}_d^i$
$G_j(\mathbf{x})$	Finite element approximation of g_j at $\mathbf{x}$
$G_i^*(\mathbf{x})$	Finite element approximation of g_i^* at $\mathbf{x}$
$e_j(\mathbf{x})$	The discretization error at $\mathbf{x}$ in the finite element approximation of g_j
$e_i^*(\mathbf{x})$	The discretization error at $\mathbf{x}$ in the finite element approximation of g_i^*
$\alpha(\mathbf{x})$	Small perturbation over the background absorption coefficient μ_a at $\mathbf{x}$
$\beta(\mathbf{x})$	Small perturbation over the background diffusion coefficient D at $\mathbf{x}$
$\Gamma_{i,j}$	Differential measurement at the i^{th} detector due to the j^{th} source
$\mathcal{A}_a$	The integral operator mapping $\alpha \in L^2(\Omega)$ to $\Gamma \in \mathbb{C}^{N_d \times N_s}$
$\mathcal{B}_b$	The integral operator mapping $\beta \in L^2(\Omega)$ to $\Gamma \in \mathbb{C}^{N_d \times N_s}$
$\mathcal{A}_a^*$	The adjoint of $\mathcal{A}_a$ mapping from $\mathbb{C}^{N_d \times N_s}$ to $L^2(\Omega)$
$\mathcal{B}_b^*$	The adjoint of $\mathcal{B}_b$ mapping from $\mathbb{C}^{N_d \times N_s}$ to $L^2(\Omega)$
$H_{i,j}^a(\mathbf{x})$	The kernel of $\mathcal{A}_a$ at $\mathbf{x}$
$H_{i,j}^b(\mathbf{x})$	The kernel of $\mathcal{B}_b$ at $\mathbf{x}$
$H_{i,j}^{a*}(\mathbf{x})$	The kernel of $\mathcal{A}_a^*$ at $\mathbf{x}$
$H_{i,j}^{b*}(\mathbf{x})$	The kernel of $\mathcal{B}_b^*$ at $\mathbf{x}$
$\gamma(\mathbf{x})$	$[\mathcal{A}_a^*\Gamma \ \mathcal{B}_b^*\Gamma]^T$ at $\mathbf{x}$
λ_a, λ_b	The regularization parameters
σ^λ	$[\alpha^\lambda \ \beta^\lambda]^T$, solution of the regularized inverse problem
$\tilde{\sigma}^\lambda$	$[\tilde{\alpha}^\lambda \ \tilde{\beta}^\lambda]^T$, solution of the regularized inverse problem with degenerate kernels
$\tilde{\sigma}_{n,m}^\lambda$	$[\tilde{\alpha}_n^\lambda \ \tilde{\beta}_m^\lambda]^T$, solution of the discretized regularized inverse problem with degenerate kernels

Table 5.2: Definition of function spaces and norms.

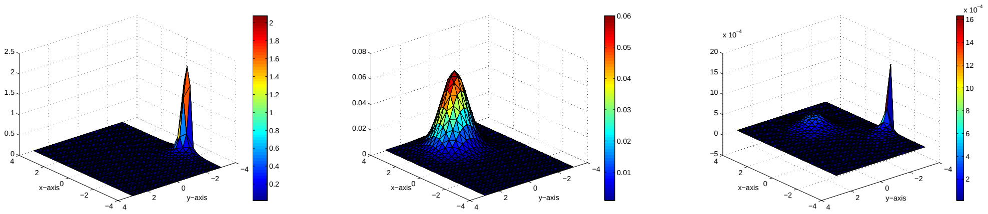
Notation	Explanation
$\overline{f}$	The complex conjugate of the function f
$C(\overline{\Omega})$	Space of continuous complex-valued functions on $\Omega \cup \partial\Omega$
$L^\infty(\Omega)$	$L^\infty(\Omega) = \{f \mid ess\,\sup_\Omega \lvert f(\mathbf{x})\rvert < \infty\ \}$
$L^p(\Omega)$	$L^p(\Omega) = \{f \mid (\int_\Omega \lvert f(\mathbf{x})\rvert^p d\mathbf{x})^{1/p} < \infty\ \}$, $p \in [1, \infty)$
$D^z_w f$	z^{th} weak derivative of f
$H^p(\Omega)$	$H^p(\Omega) = \{f \mid (\sum_{\lvert z\rvert \le p} \lVert D^z_w f\rVert^2_0)^{1/2} < \infty\ \}$, $p \in [1, \infty)$
$\lVert f\rVert_0$	The $L^2(\Omega)$ norm of f
$\lVert f\rVert_p$	The $H^p(\Omega)$ norm of f
$\lVert f\rVert_\infty$	The $L^\infty(\Omega)$ norm of f
$\lVert f\rVert_{0,m}$	The L^2 norm of f over the m^{th} finite element Ω_m
$\lVert f\rVert_{p,m}$	The H^p norm of f over the m^{th} finite element Ω_m
$\lVert f\rVert_{\infty,m}$	The L^∞ norm of f over the m^{th} finite element Ω_m



(a) A typical solution g_j for a point source located at (-1.25,-2.9).

(b) A typical reconstruction for α^λ

(c) $g_j\alpha^\lambda$.



(d) The magnitude of the gradient of the solution g_j for a point source located at (-1.25,-2.9).

(e) The absolute value of a typical reconstruction for β^λ

(f) $\lvert\nabla g_j\rvert\lvert\beta^\lambda\rvert$.

Figure 5.2: Samples of the reconstructed images in the second simulation study.

CHAPTER 6
Conclusion

In this thesis, we present an error analysis to show the relationship between the error in the reconstructed optical images and the discretization of the forward and inverse problems. We first present the error analysis which provide an insight into the impact of forward and inverse problem discretizations on the accuracy of the reconstructed optical absorption images. Next, we extend the analysis to show the relationship between the error in the simultaneously reconstructed optical absorption and diffusion coefficient images and the discretization of the forward and inverse problems.

We summarize the results of the error analysis in theorems which display the parameters that control the extent of the effect of discretization on the accuracy of optical imaging. The two most important implications of the error analysis can be stated as follows:

1. For both cases (that is optical absorption imaging and simultaneous reconstruction of optical absorption and diffusion coefficients), the analysis indicates that one has to take into account the solution of one problem (say the forward problem) while discretizing the other problem (that is the inverse problem).
2. In the case of simultaneous reconstruction of optical absorption and diffusion coefficients, the error in the reconstruction of one optical parameter (say the absorption coefficient) depends on the discretization of the other optical parameter (that is the diffusion coefficient).

The error analysis presented in this thesis motivates the development of novel adaptive discretization schemes based on the error estimates presented in chapters 2 and 5. In this respect, we developed two new adaptive mesh generation algorithms, one for the forward and one for the inverse problem in the case of optical absorption imaging. For the simultaneous reconstruction problem, the analysis lead to another set of new adaptive mesh generation algorithms, which are presented in chapter 5.

The new adaptive mesh generation algorithms take into account the interdependence between the solutions of the forward and inverse problems. Our numerical experiments in chapter 3 provided a verification of the error estimates presented in chapter 2 and showed that the proposed mesh generation algorithms significantly improve the accuracy of the reconstructed optical images for a given number of unknowns in the discrete forward and inverse problems.

Finally, we note that the error analysis and the adaptive mesh generation algorithms introduced in this thesis are not limited to DOT, and can easily be adapted for similar inverse parameter estimation problems such as electrical impedance tomography, bioluminescence tomography, optical fluorescence tomography, microwave imaging etc.

In chapter 3, we present an error analysis to show the effect of linearization of the inverse problem based on Born approximation on the accuracy of DOT image reconstructions. First, we derive an upper bound for the norm of the scattered optical field due to an absorptive perturbation. The bound is shown to depend on the spatial orientation of the optical heterogeneity and the optical field generated by the light sources. The bound provides a sensitive measure of the scattered field, since it takes the spatial variations of both the optical field and the optical heterogeneities into account. This makes the error bound significant especially in optical media with relatively small sizes. Next, we present an error analysis to show the effect of zeroth-order Tikhonov regularization and Born approximation on the accuracy of DOT imaging. The analysis shows that the error in the reconstructed image due to Born approximation depends spatially on the optical heterogeneity, the optical field generated by the light sources, and the source-detector orientation. The error analysis indicates that there is a tradeoff between the norm of the scattered field and the accuracy of the inverse problem solution. Furthermore, the error analysis provides a good measure for the choice of step length in iterative linearization based nonlinear optimization methods, such as trust-region algorithms. The error analysis presented in this work can be extended to show the effect of linearization on the accuracy of simultaneous reconstruction of scattering and absorption coefficients. Note that the presented error analysis is not limited to DOT, and can easily be adapted for similar inverse parameter estimation problems, in which Born approximation is applicable.

LITERATURE CITED

[1] M. Abramowitz and I. A. Stegun. *Handbook of Mathematical Functions.* Dover, 1968.

[2] R. Pierri ad F. Soldovieri. On the information content of the radiated fields in the near zone over bounded domains. *Inverse Problems*, 14:321–337, 1998.

[3] M. Ainsworth and J. T. Oden. A unified approach to a posteriori error estimation using elemental residual methods. *Numerische Mathematik*, 65:23–50, 1993.

[4] S. R. Arridge. Optical tomography in medical imaging. *Inverse Problems*, 15:R41–93, 1999.

[5] S. R. Arridge, J. P. Kaipio, V. Kolehmainen, M. Schweiger, E. Somersalo, T. Tarvainen, and M. Vauhkonen. Approximation errors and model reduction with an application in optical diffusion tomography. *Inverse Problems*, 22:175–195, 2006.

[6] S. R. Arridge and W. R. B. Lionheart. Non-uniqueness in diffusion-based optical tomography. *Optics Letters*, 23:882–884.

[7] S. R. Arridge and M. Schweiger. A general framework for iterative reconstruction algorithms in optical tomography, using a finite element method. *Computational radiology and imaging: theory and diagnosis*, 1998. C. Brogers and F. Natterer, eds. IMA Volumes in Mathematics and its Application Springer.

[8] K. Atkinson and H. Weimin. *Theoretical Numerical Analysis A Functional Analysis Framework.* Springer Verlag, 2001.

[9] O. Axelsson and V. Barker. *Finite Element Solution of Boundary Value Problems, Theory and Computation.* Academic Press, Orlando, FL.

[10] I. Babuška and W. C. Rheinboldt. Error estimates for adaptive finite element computations. *SIAM Journal on Numerical Analysis*, 15:736–754, 1978.

[11] I. Babuška, O. C. Zienkiewicz, J. Gago, and E. R. de A. Oliveira. *Accuracy Estimates and Adaptive Refinements in Finite Element Computations.* John Wiley and Sons, 1986.

[12] W. Bangerth. *Adaptive Finite Element Methods for the Identification of Distributed Parameters in Partial Differential Equations.* PhD thesis, University of Heidelberg, 2002.

[13] R. E. Bank and A. Weiser. Some a posterior error estimators for elliptic partial differential equations. *Mathematics of Computation*, 44:283–301, 1985.

[14] L. Beilina. *Adaptive Hybrid FEM/FDM Methods for Inverse Scattering Problems.* PhD thesis, Chalmers University of Technology, 2002.

[15] D. Boas, D. Brooks, E. Miller, C. DiMarzio, M. Kilmer, R. J. Gaudette, and Q. Zhang. Imaging the body with diffuse optical tomography. *IEEE Signal Processing Magazine*, pages 57–75, 2001.

[16] D. Boas, T. Gaudette, and S. Arridge. Simultaneous imaging and optode calibration with diffuse optical tomography. *Optics Express*, 8:263–270, 2001.

[17] D. A. Boas. A fundamental limitation of linearized algorithms for diffuse optical tomography. *Optics Express*, 1(13):404–413, 1997.

[18] A. Brancaccio, G. Leone, and R. Pierri. Information content of born scattered fields: results in the circular cylindrical case. *Journal of Optical Society of America A*, 15:1909–1917.

[19] J. G. Brankov, Y. Yang, and M. N. Wernick. Tomographic image reconstruction based on a content-adaptive mesh model. *IEEE Transactions on Medical Imaging*, 23(2):202–212, 2004.

[20] S. C. Brenner and L. R. Scott. *The Mathematical Theory of Finite Element Methods.* Springer Verlag, 2002.

[21] A. Cerussi, A. Berger, F. Bevilacqua, N. Shah, D. Jakubowski, J. Butler, R. F. Holcombe, and B. J. Tromberg. Sources of absorption and scattering contrast for near-infrared optical mammography. *Academic Radiology*, 8:211–218.

[22] B. Chance, S. Nioka, and Y. Chen. Shining new light on brain function. *OE Magazine*, 3:16–19.

[23] J. Chang, H. L. Graber, and R. L. Barbour. Imaging of fluorescence in highly scattering media. *IEEE Trans. on Biomedical Engineering*, 44:810–822, 1997.

[24] X. Cheng and D. Boas. Systematic diffuse optical image errors resulting from uncertainty in the background optical properties. *Optics Express*, 4(8):299–307, 1999.

[25] D. Colton and R. Kress. *Inverse Acoustic and Electromagnetic Scattering Theory*. Springer-Verlag, 1998.

[26] A. Corlu, T. Durduran, Re. Choe, M. Schweiger, E. M. C. Hillman, S. R. Arridge, and A. G. Yodh. Uniqueness and wavelength optimization in continuous-wave multispectral diffuse optical tomography. *Optics Letters*, 28:2339–2341.

[27] D. Daners. Robin boundary value problems on arbitrary domains. *Transactions of the American Mathematical Society*, 352(9):4207–4236, 2000.

[28] M. J. Eppstein, D. E. Dougherty, D. J. Hawrysz, and E. M. Sevick-Muraca. Three-dimensional Bayesian optical image reconstruction with domain decomposition. *IEEE Trans. on Medical Imaging*, 20(3):147–163, 2001.

[29] M. J. Eppstein, D. J. Hawrysz, A. Godavarty, and E. M. Sevick-Muraca. Three dimensional near infrared fluorescence tomography with bayesian methodologies for image reconstruction from sparse and noisy data sets. volume 99 of *Proceedings of National Academy of Science*, pages 9619–9624.

[30] L. C. Evans. *Partial Differential Equations*. American Mathematical Society, 1998.

[31] R. Eymard, T. Gallouët, and R. Herbin. *Finite volume methods*, volume VII of *Handbook of numerical analysis.* 2000.

[32] J. V. Frangioni. In vivo near-infrared fluorescence imaging. *The Journal of Sports Medicine and Physical Fitness*, 7:626–634.

[33] R. J. Gaudette, D. H. Brooks, C. A. DiMarzio, M. E. Kilmer, E. L. Miller, T. Gaudette, and D. A. Boas. A comparison study of linear reconstruction techniques for diffuse optical tomographic imaging of absorption coefficient. *Physics in Medicine and Biology*, 45:1051–1070, 2000.

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

[35] X. Gu, Y. Xu, and H. Jiang. Mesh-based enhancement schemes in diffuse optical tomography. *Medical Physics*, 30(5):861–869, 2003.

[36] M. Guven, B. Yazici, X. Intes, and B. Chance. An adaptive multigrid algorithm for region of interest diffuse optical tomography. In *International Conference in Image Processing*, volume 2 of *Proc. of IEEE*, pages 823–826, 2003.

[37] M. Guven, B. Yazici, X. Intes, and B. Chance. Diffuse optical tomography with a priori anatomical information. *Physics in Medicine and Biology*, 50:2837–2858, 2005.

[38] M. Guven, B. Yazici, K. Kwon, E. Giladi, and X. Intes. Error in optical absorption images due to Born approximation in diffuse optical tomography. in preparation.

[39] M. Guven, B. Yazici, K. Kwon, E. Giladi, and X. Intes. Effect of discretization error and adaptive mesh generation in diffuse optical absorption imaging: I. *Inverse Problems*, 23:1115–1133, 2007.

[40] M. Guven, B. Yazici, K. Kwon, E. Giladi, and X. Intes. Effect of discretization error and adaptive mesh generation in diffuse optical absorption imaging: II. *Inverse Problems*, 23:1135–1160, 2007.

[41] D. J. Hawrysz and E. M. Sevick-Muraca. Developments toward diagnostic breast cancer imaging using near-infrared optical measurements and fluorescent contrast agents. *Neoplasia*, 2(5):388–417, 2000.

[42] J. C. Hebden and S. R. Arridge D. T. Delpy. Optical imaging in medicine: I. Experimental techniques. *Physics in Medicine and Biology*, 42:825–840, 1997.

[43] J. C. Hebden, A. Gibson, T. Austin, R. Yusof, N. Everdell, E. Delpy, S. R. Arridge, J. H. Meek, and J. S. Wyatt. Imaging changes in blood volume and oxygenation in the newborn infant brain using three-dimensional optical tomography. *Physics in Medicine and Biology*, 49:1117–1130.

[44] J. C. Hebden, A. Gibson, R. Md Yusof, N. Everdell, E. M. C. Hillman, D. T. Delpy, S. R. Arridge, T. Austin, J. H. Meek, and J. S. Wyatt. Three-dimensional optical tomography of the premature infant brain. *Physics in Medicine and Biology*, 47:4155–4166.

[45] V. Heuveline and R. Rannacher. Duality-based adaptivity in the hp-finite element method. *Journal of Numerical Mathematics*, 11:95–113.

[46] M. Huang and Q. Zhu. Dual-mesh optical tomography reconstruction method with a depth correction that uses a priori ultrasound information. *Applied Optics*, 43(8):1654–1662, 2004.

[47] F. Ihlenburg. *Finite Element Analysis of Acoustic Scattering.* Springer Verlag, 1998.

[48] X. Intes. Projection access order in algebraic reconstruction technique for diffuse optical tomography. *Physics in Medicine and Biology*, 47:N1–N10, 2002.

[49] X. Intes and B. Chance. Non-PET functional imaging techniques: Optical. *The Radiologic Clinics of North America*, 43:221–234, 2005.

[50] X. Intes, C. Maloux, M. Guven, B. Yazıcı, and B. Chance. Diffuse optical tomography with physiological and spatial a priori constraints. *Physics in Medicine and Biology*, 49:N155–63, 2004.

[51] X. Intes, J. Ripoll, Y. Chen, S. Nioka, A. Yodh, and B. Chance. In vivo continuous-wave optical breast imaging enhanced with indocyanine green. *Medical Physics*, 30:1039–1047.

[52] V. Isakov. On uniqueness in the inverse transmission scattering problem. *Commun. Partial Diff. Eq.*, 15:1565–87, 1990.

[53] J. D. Jackson. *Classical Electrodynamics.* John Wiley and Sons, 1962.

[54] A. Joshi, W. Bangerth, and E. M. Sevick-Muraca. Adaptive finite element based tomography for fluorescence optical imaging in tissue. *Optics Express*, 12(22):5402–5417, 2004.

[55] J. Kaipio and E. Somersalo. *Statistical and computational inverse problems*, volume 160 of *Applied Mathematical Sciences.* Springer-Verlag, New York, 2005.

[56] K. Kincade. Optical diagnostics continue migration from bench top to beside. *Laser Focus World*, pages 130–134, 2004.

[57] R. Kress. *Linear integral equations*, volume 82 of *Applied Mathematical Sciences.* Springer-Verlag, second edition, 1999.

[58] I. W. Kwee. *Towards a Bayesian Framework for Optical Tomography.* PhD thesis, University of London, 1999.

[59] K. Kwon and B. Yazıcı. Born expansion and Fréchet derivatives in diffuse optical tomography. In preparation.

[60] K. Kwon, B. Yazıcı, and M. Guven. Two-level domain decomposition methods for diffuse optical tomography. *Inverse Problems*, 22:1533–1559, 2006.

[61] A. Li, G. Boverman, Y. Zhang, D. Brooks, E. L. Miller, M. E. Kilmer, Q. Zhang, E. M. C. Hillman, and D. A. Boas. Optimal linear inverse solution with multiple priors in diffuse optical tomography. *Applied Optics*, 44:1948–1956, 2005.

[62] R. Li, W. Liu, H. Ma, and T. Tang. Adaptive finite element approximation for distributed elliptic optimal control problems. *SIAM Journal on Control and Optimization*, 41:1321–1349, 2002.

[63] V. A. Markel, V. Mital, and J. C. Schotland. Inverse problem in optical diffusion tomography. III. inversion formulas and singular-value decomposition. *Journal of Optical Society of America A*, 20(5):890–902, 2003.

[64] V. A. Markel and J. C. Schotland. Inverse problem in optical diffusion tomography. II. role of boundary conditions. *Journal of Optical Society of America A*, 19(3):1–9, 2002.

[65] T. O. McBride, B. W. Pogue, U. L. sterberg, and K. D. Paulsen. Separation of absorption and scattering heterogeneities in nir tomographic imaging of tissue. Biomedical Topical Meetings, pages 339–341. OSA Technical Digest, 2000.

[66] M. Molinari, S. J. Cox, B. H. Blott, and G. J. Daniell. Adaptive mesh refinement techniques for electrical impedance tomography. *Physiological Measurement*, 22:91–96, 2001.

[67] F. Natterer. An error bound for the Born approximation. *Inverse Problems*, 20:447–452, 2004.

[68] V. Ntziachristos, B. Chance, and A. G. Yodh. Differential diffuse optical tomography. *Opt. Exp.*, 5:230–242, 1999.

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

[70] V. Ntziachristos, A. G. Yodh, M. D. Schnall, and B. Chance. Mri-guided diffuse optical spectroscopy of malignant and benign breast lesions. *Neoplasia*, 4:347–354, 2002.

[71] J. T. Oden and S. Prudhomme. Goal-oriented error estimation and adaptivity for the finite element method. *Computers and Mathematical Applications*, 41:735–756.

[72] M. A. O'leary. *Imaging with Diffuse Photon Density Waves.* PhD thesis, University of Pennsylvania, 1999.

[73] G. Y. Panasyuk, V. A. Markel, and J. C. Schotland. Superresolution and corrections to the diffusion approximation in optical tomography. *Applied Physics Letters*, 87:1–3.

[74] Y. Pei, H. Graber, and R. Barbour. Normalized-constraint algorithm for minimizing inter-parameter crosstalk in dc optical tomography. *Optics Express*, 9:97–109.

[75] V. Quaresima, R. Lepanto, and M. Ferrari. The use of near infrared spectroscopy in sports medicine. *The Journal of Sports Medicine and Physical Fitness*, 43:1–13.

[76] M. C. Rivara. Mesh refinement processes based on the generalized bisection of simplices. *SIAM Journal on Numerical Analysis*, 21(3):604–613, 1984.

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

[78] G. D. Smith. *Numerical Solution of Partial Differential Equations: Finite Difference Methods.* Oxford Applied Math. Comput. Sci. Series. Oxford: Clarendon, third edition, 1985.

[79] A. Srinivasan, B. Pogue, S. Jiang, H. Dehghani, C. Kogel, S. Soho, J. J. Gibson, T. D. Tosteson, S. P. Poplack, and K. Paulsen. Interpreting hemoglobin and water concentration, oxygen saturation, and scattering measured in vivo by near-infrared breast tomography. volume 100 of *Proceedings of National Academy of Science*, pages 12349–12354.

[80] G. Strangman, D. Boas, and J. Sutton. Non-invasive neuroimaging using near-infrared light. *Biological Psychiatry Psychiatry*, 52:679–693.

[81] T. Strouboulis and K. A. Hague. Recent experiences with error estimation and adaptivity, part I: Review of error estimators for scalar elliptic problems. *Computer Methods in Applied Mechanics and Engineering*, 97:399–436, 1992.

[82] M. Torregrossa, C. V. Zint, and P. Poulet. Image reconstruction in optical tomography: mesh influence. In *IV International Workshop*, Computational Problems of Electrical Engineering, pages 183–186, 2002.

[83] R. Verfurth. *A Review of A Posteriori Error Estimation and Adaptive Mesh Refinement Techniques*. Teubner-Wiley, 1996.

[84] B. Villringer and B. Chance. Non-invasive optical spectroscopy and imaging of human brain function. *Trends in Neuroscience*, 20:435–442.

[85] S. Walker, D. A. Boas, and E. Gratton. Photon density waves scattered from cylindrical inhomogeneities: theory and experiments. *Applied Optics*, 37(10):1935–1944, 1998.

[86] R. Weissleder and V. Ntziachristos. Shedding light onto live molecular targets. *Nature Medicine*, 9:123–128.

[87] A. G. Yodh and B. Chance. Spectroscopy and imaging with diffusing light. *Physics Today*, 48:34–40, 1995.

APPENDIX A
Appendix for chapter 2

A.1 Boundedness of $\mathcal{A}_a$

$$\|\mathcal{A}_a \alpha\|_{l^1} = \sum_{i,j}^{N_d,N_s} \left| \int_\Omega H_{i,j}(\mathbf{x}) \alpha(\mathbf{x}) d\mathbf{x} \right|. \tag{A.1}$$

We can write the following inequality:

$$\begin{aligned} \|\mathcal{A}_a \alpha\|_{l^1} &\leq \sum_{i,j}^{N_d,N_s} \int_\Omega |H_{i,j}(\mathbf{x}) \alpha(\mathbf{x})| d\mathbf{x} \\ &\leq \left(\sum_{i,j}^{N_d,N_s} \int_\Omega |H_{i,j}(\mathbf{x})| d\mathbf{x} \right) \|\alpha\|_\infty. \end{aligned} \tag{A.2}$$

Using Schwarz' inequality, we can write an upper bound for the summation as follows:

$$\begin{aligned} \sum_{i,j}^{N_d,N_s} \int_\Omega |H_{i,j}(\mathbf{x})| d\mathbf{x} &= \sum_{i,j}^{N_d,N_s} \|g_i^* g_j\|_{L^1(\Omega)} \\ &\leq \sum_{i,j}^{N_d,N_s} \|g_i^*\|_0 \|g_j\|_0 \\ &\leq N_d N_s \max_i \|g_i^*\|_0 \max_j \|g_j\|_0, \end{aligned} \tag{A.3}$$

which leads to

$$\|\mathcal{A}_a \alpha\|_{l^1} \leq N_d N_s \max_i \|g_i^*\|_0 \max_j \|g_j\|_0 \|\alpha\|_\infty.$$

Therefore an upper bound for the norm of $\mathcal{A}_a$ is given by

$$\|\mathcal{A}_a\|_{L^\infty(\Omega) \to l^1} \leq N_d N_s \max_i \|g_i^*\|_0 \max_j \|g_j\|_0. \tag{A.4}$$

The boundedness of g_j and g_i^* imply that $\mathcal{A}_a$ is bounded. □

A.2 Compactness of $\mathcal{A}_a$

$\mathcal{A}_a$ is bounded by (A.4). Furthermore $\mathcal{A}_a$ maps the infinite dimensional subspace $L^\infty(\Omega)$ to a finite dimensional subspace $\mathbb{C}^{N_d \times N_s}$, that is the range $R(\mathcal{A}_a)$ of $\mathcal{A}_a$ satisfies $R(\mathcal{A}_a) \in \mathbb{C}^{N_d \times N_s}$ due to the finite number of sources and detectors. As a result, $\mathcal{A}_a$ is compact [57]. The inverse problem is ill-posed as a consequence of compactness [57]. □

A.3 Proof of the Lemma

The identity operator $\mathcal{I}$ is a bounded operator with bounded inverse and $(\mathcal{P}_n\mathcal{I})^{-1} = \mathcal{I} : X_n \to X_n$. Furthermore, $\|\mathcal{P}_n\|_{X \to X_n}$ is bounded for first order Lagrange basis functions [8, 57]. Thus, projection by collocation converges for the identity operator. $\mathcal{A}$ is bounded and compact, and $\mathcal{K} = \lambda\mathcal{I} + \mathcal{A}$ is injective, with bounded inverse given by (4.33). As a result, by Theorem 13.7 in [57], the projection method also converges for $\mathcal{K} = \lambda\mathcal{I} + \mathcal{A}$. Convergence of projection for $\mathcal{K}$ implies $(\mathcal{P}_n\mathcal{K})^{-1}\mathcal{P}_n\mathcal{K}\alpha^\lambda \to \alpha^\lambda$, $n \to \infty$ for $(\mathcal{P}_n\mathcal{K})^{-1}\mathcal{P}_n\mathcal{K} : X \to X_n$ [57].

It follows from the proof of Theorem 13.7 in [57] that $(\mathcal{I} + \frac{1}{\lambda}\mathcal{P}_n\mathcal{A})^{-1} : Y_n \to X_n$ exists and is uniformly bounded for all sufficiently large n. Then from $\mathcal{P}_n\mathcal{K} = \lambda\mathcal{P}_n(\mathcal{I} + \frac{1}{\lambda}\mathcal{P}_n\mathcal{A}) = \lambda(\mathcal{I} + \frac{1}{\lambda}\mathcal{P}_n\mathcal{A})$, it follows that $\mathcal{P}_n\mathcal{K} : X_n \to Y_n$ is invertible for all sufficiently large n with the inverse given by

$$(\mathcal{P}_n\mathcal{K})^{-1} = (\mathcal{I} + \frac{1}{\lambda}\mathcal{P}_n\mathcal{A})^{-1}\frac{1}{\lambda}. \tag{A.5}$$

As a result we can write $(\mathcal{P}_n\mathcal{K})^{-1}\mathcal{P}_n\mathcal{K}$ as follows:

$$(\mathcal{P}_n\mathcal{K})^{-1}\mathcal{P}_n\mathcal{K} = (\mathcal{I} + \frac{1}{\lambda}\mathcal{P}_n\mathcal{A})^{-1}\frac{1}{\lambda}\mathcal{P}_n\mathcal{K}. \tag{A.6}$$

Thus,

$$\|(\mathcal{P}_n\mathcal{K})^{-1}\mathcal{P}_n\mathcal{K}\|_{X \to X_n} \leq C_M \frac{\|\mathcal{K}\|_{X \to Y}}{\lambda} \tag{A.7}$$

where $C_M > 0$ is independent of n, using the facts that projection by collocation method converges for the identity operator and $(\mathcal{I} + \frac{1}{\lambda}\mathcal{P}_n\mathcal{A})^{-1}$ is uniformly bounded.

□

APPENDIX B

Appendix for chapter 3

B.1 Solution of the Model Problem (shorter version)

In order to initialize the adaptive mesh for the solution of the forward problem (provided $D(\mathbf{x}) = \overline{D}$ and $\mu_a(\mathbf{x}) = \overline{\mu_a}$ are spatially constant), we use an analytical solution to compute the estimates of g_j and g_i^*. Below, we give the solution in 2D for (5.1). Under the same conditions, an analytical solution for the adjoint problem (5.3) can be obtained in a similar way.

First, we use the polar coordinates (ρ,θ) to rewrite (5.1):

$$\frac{1}{\rho}\frac{\partial}{\partial\rho}(\rho\frac{\partial g_j}{\partial\rho}) + \frac{1}{\rho}\frac{\partial}{\partial\theta}(\rho\frac{\partial g_j}{\partial\theta}) + K_\Omega^2 g_j \quad = \quad -\frac{4\pi}{\rho}\frac{\delta(\rho-\rho_s^j)\delta(\theta-\theta_s^j)}{\overline{D}},$$

where we consider an unbounded domain, model the point source located at (ρ_s^j,θ_s^j) by the Dirac-delta function $4\pi\delta(\rho-\rho_s^j)\delta(\theta-\theta_s^j)/\rho$, and $K_\Omega^2 = -(\overline{\mu_a}c + i\omega)/c\overline{D}$. Then, the solution g_j at (ρ,θ) due to the point source located at (ρ_s^j,θ_s^j) is given by [53]

$$g_j(\rho,\rho_s^j;\theta,\theta_s^j) = \frac{4}{\overline{D}\pi}\left\{\ \tfrac{1}{2}I_0(k_\Omega\rho_<)K_0(k_\Omega\rho_>) + \textstyle\sum_{m=1}^{\infty}\cos[m(\theta-\theta_s)]I_m(k_\Omega\rho_<)K_m(k_\Omega\rho_>)\ \right\},$$

where $\rho_<$ means the smaller of ρ and ρ_s^j, $\rho_>$ means the greater of ρ and ρ_s^j, I_m and K_m are the modified Bessel functions of the first and second kind, respectively [1] and $k_\Omega = \sqrt{-K_\Omega^2}$.

The solution of the problem in 3D can be derived in a similar manner [53, 85].

B.2 Solution of the Model Problem

Consider a circular heterogeneity in an unbounded medium. Figure B.2 shows such a heterogeneity. In this section, we solve the diffusion equation on such a domain, in which the optical properties of the circular inclusion and the background are different. For this purpose, we first derive a solution for a homogeneous optical medium. Next, we update our solution to account for the scattered field originating from the cylindrical heterogeneity.

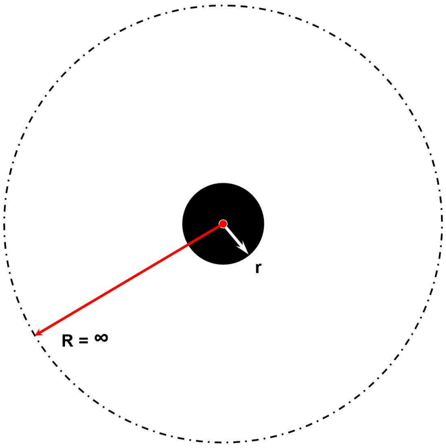


Figure B.1: The setup considered for the solution of the model problem in equation (B.1).

B.2.1 Homogeneous Optical Background

First, consider the constant-coefficient frequency-domain diffusion equation in $3D$:

$$(\nabla^2 + K_{(bg)}^2)\Phi_{inc} = \frac{-\delta(x-x_s)\delta(y-y_s)\delta(z-z_s)}{D_{(bg)}}, \tag{B.1}$$

where Φ_{inc} is the incident optical density, $K_{(bg)}^2 = -(\mu_{a(bg)}c + j\omega)/cD_{(bg)}$, $\mu_{a(bg)}$ and $D_{(bg)}$ are the absorption and diffusion coefficients of the background, respectively.

For our model problem, we consider a circular heterogeneity in an infinite homogeneous medium, hence cylindrical (polar) coordinate system is appropriate to expand the solution of (B.1). Therefore (B.1) can be rewritten in cylindrical coordinates as follows:

$$\frac{1}{\rho}\frac{\partial}{\partial\rho}(\rho\frac{\partial\Phi_{inc}}{\partial\rho}) + \frac{1}{\rho}\frac{\partial}{\partial\theta}(\rho\frac{\partial\Phi_{inc}}{\partial\theta}) + \frac{\partial^2\Phi_{inc}}{\partial z^2} + K_{(bg)}^2\Phi_{inc} = -\frac{4\pi}{\rho}\frac{\delta(\rho-\rho_s)\delta(\theta-\theta_s)\delta(z-z_s)}{D_{(bg)}}. \tag{B.2}$$

We make use of the following identities for the solution of equation (B.2) [53]:

$$\begin{aligned}\delta(z-z_s) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} ds e^{is(z-z_s)} = \frac{1}{\pi}\int_0^{\infty} ds\cos[s(z-z_s)], \\ \delta(\theta-\theta_s) &= \frac{1}{2\pi}\sum_{m=-\infty}^{\infty} e^{im(\theta-\theta_s)},\end{aligned} \tag{B.3}$$

and we expand the solution in cylindrical coordinates as

$$\Phi_{inc}(\mathbf{x}, z; x_s, y_s, z_s) = \frac{1}{2\pi^2}\sum_{m=-\infty}^{\infty}\int_0^{\infty} ds e^{im(\theta-\theta_s)} cos[s(z-z_s)] g_m(k,\rho,\rho_s), \tag{B.4}$$

where $g_m(k,\rho,\rho_s)$ is the radial component of the solution Φ_{inc}, as a function of m, k, ρ and ρ_s.

Substituting equations (B.3) and (B.4) in equation (B.2) and rearranging the terms yields the following differential equation:

$$\frac{d^2}{d\rho^2}g_m + \frac{2}{\rho}\frac{d}{d\rho}g_m - (s^2 + \frac{m^2}{\rho^2} - K_{(bg)}^2)g_m = -\frac{4\pi}{D_{(bg)}\rho}\delta(\rho-\rho_s). \tag{B.5}$$

Equation B.5 is similar to the modified Bessel equation for $\rho \neq \rho_s$. In this case, the substitution $\sqrt{s^2 - K_{(bg)}^2}\rho = \tau\rho = r$ yields

$$\frac{d^2}{dr^2}g_m + \frac{1}{r}\frac{d}{dr}g_m - (1 + \frac{m^2}{r^2})g_m = 0. \tag{B.6}$$

Thus, for $\rho \neq \rho_s$, the solution to equation B.5 is given by the modified Bessel functions, $I_m(\tau\rho)$ and $K_m(\tau\rho)$.

We assume that the following conditions hold for the solution:

1. $g_m(k,\rho,\rho_s)$ be finite at $\rho = 0$,
2. $g_m(k,\rho,\rho_s)$ vanishes at $\rho \to \infty$.

Suppose that $\psi_1(\tau\rho)$ is some linear combination of I_m and K_m which satisfies the correct boundary conditions for $\rho < \rho_s$ and that $\psi_2(\tau\rho)$ is a linearly independent combination which satisfies the proper boundary conditions for $\rho > \rho_s$. The symmetry of the solution [53] in ρ and ρ_s, that is $g_m(k,\rho,\rho_s) = g_m(k,\rho_s,\rho)$, requires that

$$g_m(\rho, \rho_s) = \psi_1(\tau\rho_<)\psi_2(\tau\rho_>) \tag{B.7}$$

where $\rho_<$ means smaller of ρ and ρ_s and $\rho_>$ implies larger of ρ and ρ_s. $\psi_1(\tau\rho)$ is the solution for $\rho < \rho_s$, thus the finiteness requirement at $\rho = 0$ implies the coefficient B_1 should be zero (since the modified Bessel function $K_m \to \infty$ as $\rho \to 0$) (Abramowitz and Stegun 1968). Similarly, the second requirement implies the coefficient A_2 in $\psi_2(\tau\rho)$ is equal to zero, since $\psi_2(\tau\rho)$ is the solution for $\rho > \rho_s$ and $I_m \to \infty$ as $\rho \to \infty$. Thus, we conclude that the solution $g_m(k, \rho, \rho_s)$ is given by $g_m(k, \rho, \rho_s) = A\psi_1(\tau\rho)\psi_2(\tau\rho)$, for some constant A. The constant A is to be determined from the Wronskian condition, which results in $A = \frac{4\pi}{D_{(bg)}}$ [53]. Then, the incident field $\Phi_{inc}(\rho, \theta, z; \rho_s, \theta_s, z_s)$ in 3D is given by

$$\Phi_{inc}(\rho, \theta, z; \rho_s, \theta_s, z_s) = \frac{2}{D_{(bg)}\pi} \sum_{m=-\infty}^{\infty} \int_0^{\infty} ds e^{im(\theta-\theta_s)} cos[s(z - z_s)] I_m(\tau\rho_<) K_m(\tau\rho_>). \tag{B.8}$$

Noting that $I_n = I_{-n}$ and $K_n = K_{-n}$ for n an integer [1], equation (B.8) can be equivalently written in terms of real functions as follows:

$$\begin{aligned}\Phi_{inc}(\rho, \rho_s) = \frac{4}{D_{(bg)}\pi} \int_0^{\infty} ds \cos[s(z - z_s)] \{ \frac{1}{2} I_0(\tau\rho_<) K_0(\tau\rho_>) \\ + \sum_{m=1}^{\infty} \cos[m(\theta - \theta_s)] I_m(\tau\rho_<) K_m(\tau\rho_>) \}\end{aligned} \tag{B.9}$$

In order to compute a solution in 2D, we can integrate the expression in B.9 with respect to $(z - z_s)$ between the limits $\pm Z$, where Z is taken to be very large. An alternative way is to follow an approach similar to the one we followed in 3D case. Instead, letting $s \to 0$ and removing the related integral associated with $\delta(z - z_s)$ will yield the desired result in 2D. In this case, the solution in $2D$ is given by

$$\begin{aligned}\Phi_{inc}(\rho, \rho_s; \theta, \theta_s) &= \frac{4}{D_{(bg)}\pi} \{ \frac{1}{2} I_0(k_{(bg)}\rho_<) K_0(k_{(bg)}\rho_>) \\ &+ \sum_{m=1}^{\infty} \cos[m(\theta - \theta_s)] I_m(k_{(bg)}\rho_<) K_m(k_{(bg)}\rho_>) \},\end{aligned} \tag{B.10}$$

where $k_{(bg)} = \sqrt{-K^2_{(bg)}}$.

B.2.2 The scattered field due to the circular inclusion

Now consider a circular inclusion of radius $\rho = a$ in the homogenous background with its center at the origin (that is $\rho = 0$). When the incident wave reaches the cylinder, the wave front is distorted. We assume that the wave field outside the cylinder is a superposition of the incident field and a scattered field (due to the existence of the circular inclusion), that is $\Phi_{(out)} = \Phi_{(inc)} + \Phi_{(sc)}$. In this representation, $\Phi_{(sc)}$ and the field inside the circle $\Phi_{(in)}$ are given by the modified Bessel functions, as it should satisfy B.2 outside the circle.

The scattered field Φ_{sc} and the incident field Φ_{inc} in the presence of a circular inclusion are given in terms of the modified Bessel functions as follows:

$$\begin{aligned}\Phi_{(sc)} &= \sum_{n=0}^{\infty} \cos[n(\theta - \theta_s)] A_n^{(sc)} I_n(k_{(bg)}\rho) + B_n^{(sc)} K_n(k_{(bg)}\rho), &\text{(B.11)}\\ \Phi_{(in)} &= \sum_{n=0}^{\infty} \cos[n(\theta - \theta_s)] A_n^{in} I_n(k_{(in)}\rho) + B_n^{in} I_n(k_{(in)}\rho). &\text{(B.12)}\end{aligned}$$

The unknown coefficients in Φ_{sc} and Φ_{in} representations are to be found by applying appropriate boundary conditions at the interface of the circular inclusion and for $\rho = 0$ and $\rho \to \infty$ [85]:

1. $\Phi_{sc} \to 0$ as $\rho \to \infty$.

2. Φ_{in} is finite everywhere.

3. $(1 - R_{21})\Phi_{out} = [(1 - R_{21}) + 2(R_{12} - R_{21})D_{(in)}\frac{\partial}{\partial\rho}]\Phi_{in}$ at $\rho = a$.

4. $D_{(bg)}\frac{\partial}{\partial\rho}\Phi_{out} = D_{(in)}\frac{\partial}{\partial\rho}\Phi_{in}$ at $\rho = a$.

where R_{12} and R_{21} are the reflection coefficients at the interface [85]. The first two conditions suggest that $A_n^{(sc)}$ (since $I_n(x) \to \infty$ as $x \to \infty$) and $B_n^{(in)}$ (since $K_n(x) \to \infty$ as $x \to 0$) are equal to zero. The remaining coefficients, $A_{(in)}$ and $B_{(sc)}$ can be found using the third and fourth conditions.

$$B_n^{(sc)} = \frac{X_n V_n - Y_n}{Z_n - S_n X_n}, \tag{B.13}$$

where

$$\begin{aligned} X_n &= D_{(in)} k_{(in)} I_n'(k_{(in)} a), && \text{(B.14)} \\ Z_n &= D_{(out)} k_{(out)} K_n'(k_{(out)} a), && \text{(B.15)} \\ F_n &= (1 - R_{(21)}) I_n(k_{(in)} a) + 2(R_{12} - R_{21}) D_{(in)} k_{(in)} I_n'(k_{(in)} a)) && \text{(B.16)} \end{aligned}$$

and for $n = 0$,

$$\begin{aligned} V_0 &= \frac{2}{\pi D_{(out)}} I_0(k_{(out)} a) K_0(k_{(out)} \rho_s), && \text{(B.17)} \\ Y_0 &= \frac{2k_{(out)}}{D_{(out)}} K_0(k_{(out)} \rho_s) I_0'(k_{(out)} a), && \text{(B.18)} \end{aligned}$$

and for $n \geq 1$

$$\begin{aligned} V_n &= \frac{4}{\pi D_{(out)}} I_n(k_{(out)} a) K_0(k_{(out)} \rho_s), && \text{(B.19)} \\ Y_n &= \frac{4k_{(out)}}{D_{(out)}} K_n(k_{(out)} \rho_s) I_n'(k_{(out)} a). && \text{(B.20)} \end{aligned}$$

As a result, the coefficients $A_n^{(in)}$ are given as

$$A_n^{(in)} = \frac{S_n B_n^{(sc)} + V_n}{F_n}, \tag{B.21}$$

where

$$S_n = (1 - R_{21}) K_n(k_{(out)} a) \tag{B.22}$$

where we assumed I_n' and K_n' are the derivatives of I_n and K_n, respectively.

APPENDIX C

Appendix for chapter 4

C.1 Proof of theorem 1

We can write [39]

$$\sigma^\lambda - \tilde{\sigma}^\lambda = \tilde{\mathcal{T}}^{-1}\left\{\left(\mathcal{T} - \tilde{\mathcal{T}}\right)\sigma^\lambda + (\gamma - \tilde{\gamma})\right\}. \tag{C.1}$$

In the following, we will provide appropriate approximations for $\left(\mathcal{T} - \tilde{\mathcal{T}}\right)\sigma^\lambda$ and $(\gamma - \tilde{\gamma})$.

By definition,

$$\|(\mathcal{A}_a - \tilde{\mathcal{A}}_a)\alpha\|_{l^2}^2 = \sum_{i,j}^{N_d,N_s}\left|\int_\Omega \left(\overline{g_i^*(\mathbf{x})}g_j(\mathbf{x}) - \overline{G_i^*(\mathbf{x})}G_j(\mathbf{x})\right)\alpha(\mathbf{x})d\mathbf{x}\right|^2. \tag{C.2}$$

Similarly,

$$\|(\mathcal{B}_b - \tilde{\mathcal{B}}_b)\beta\|_{l^2}^2 = \sum_{i,j}^{N_d,N_s}\left|\int_\Omega \left(\nabla\overline{g_i^*(\mathbf{x})}\cdot\nabla g_j(\mathbf{x}) - \nabla\overline{G_i^*(\mathbf{x})}\cdot\nabla G_j(\mathbf{x})\right)\beta(\mathbf{x})d\mathbf{x}\right|^2. \tag{C.3}$$

where G_i^*, G_j are the finite element approximations to g_i^* and g_j, respectively. We can expand $\overline{g_i^*}g_j - \overline{G_i^*}G_j$ as

$$\overline{g_i^*}g_j - \overline{G_i^*}G_j = \overline{e_i^*}e_j + G_j\overline{e_i^*} + \overline{G_i^*}e_j, \tag{C.4}$$

where $e_i^* := g_i^* - G_i^*$ and $e_j := g_j - G_j$. Replacing G_i^* and G_j respectively with $g_i^* - e_i^*$ and $g_j - e_j$, we get

$$\begin{aligned}\overline{g_i^*}g_j - \overline{G_i^*}G_j &= g_j\overline{e_i^*} + \overline{g_i^*}e_j - \overline{e_i^*}e_j \\ &\approx g_j\overline{e_i^*} + \overline{g_i^*}e_j,\end{aligned} \tag{C.5}$$

where we neglect the term $\overline{e_i^*}e_j$. In a similar way, we can write

$$\begin{aligned}\nabla\overline{g_i^*}\cdot\nabla g_j-\nabla\overline{G_i^*}\cdot\nabla G_j &= \nabla g_j\cdot\nabla\overline{e_i^*}+\nabla\overline{g_i^*}\cdot\nabla e_j-\nabla\overline{e_i^*}\cdot\nabla e_j\\ &\approx \nabla g_j\cdot\nabla\overline{e_i^*}+\nabla\overline{g_i^*}\cdot\nabla e_j, \qquad (C.6)\end{aligned}$$

We can express $\mathcal{A}-\tilde{\mathcal{A}}$ as

$$\mathcal{A}-\tilde{\mathcal{A}}=\mathcal{A}_a^*\mathcal{A}_a-\tilde{\mathcal{A}}_a^*\tilde{\mathcal{A}}_a. \qquad (C.7)$$

Following a similar approach as above,

$$\mathcal{A}_a^*\mathcal{A}_a-\tilde{\mathcal{A}}_a^*\tilde{\mathcal{A}}_a=(\mathcal{A}_a^*-\tilde{\mathcal{A}}_a^*)(\mathcal{A}_a-\tilde{\mathcal{A}}_a)+\tilde{\mathcal{A}}_a^*(\mathcal{A}_a-\tilde{\mathcal{A}}_a)+(\mathcal{A}_a^*-\tilde{\mathcal{A}}_a^*)\tilde{\mathcal{A}}_a. \qquad (C.8)$$

As a result, the following condition holds:

$$\|\mathcal{A}-\tilde{\mathcal{A}}\|\leq\|(\mathcal{A}_a^*-\tilde{\mathcal{A}}_a^*)(\mathcal{A}_a-\tilde{\mathcal{A}}_a)\|+\|\tilde{\mathcal{A}}_a^*(\mathcal{A}_a-\tilde{\mathcal{A}}_a)+(\mathcal{A}_a^*-\tilde{\mathcal{A}}_a^*)\tilde{\mathcal{A}}_a\|. \qquad (C.9)$$

Since $\tilde{\mathcal{A}}_a=-(\mathcal{A}_a-\tilde{\mathcal{A}}_a)+\mathcal{A}_a$, (C.9) can be rewritten as

$$\begin{aligned}\|\mathcal{A}-\tilde{\mathcal{A}}\| &= \|\mathcal{A}_a^*\mathcal{A}_a-\tilde{\mathcal{A}}_a^*\tilde{\mathcal{A}}_a\|\\ &\leq \|(\mathcal{A}_a^*-\tilde{\mathcal{A}}_a^*)(\mathcal{A}_a-\tilde{\mathcal{A}}_a)\|+2\|\mathcal{A}_a^*(\mathcal{A}_a-\tilde{\mathcal{A}}_a)\|\\ &\approx 2\|\mathcal{A}_a^*(\mathcal{A}_a-\tilde{\mathcal{A}}_a)\|, \qquad (C.10)\end{aligned}$$

where we neglect the term $\|(\mathcal{A}_a^*-\tilde{\mathcal{A}}_a^*)(\mathcal{A}_a-\tilde{\mathcal{A}}_a)\|$. Similarly,

$$\begin{aligned}\|\mathcal{B}-\tilde{\mathcal{B}}\| &\leq 2\|\mathcal{B}_b^*(\mathcal{B}_b-\tilde{\mathcal{B}}_b)\|, \qquad (C.11)\\ \|\mathcal{A}_B-\tilde{\mathcal{A}}_B\| &\leq 2\|\mathcal{A}_a^*(\mathcal{B}_b-\tilde{\mathcal{B}}_b)\|, \qquad (C.12)\\ \|\mathcal{B}_A-\tilde{\mathcal{B}}_A\| &\leq 2\|\mathcal{B}_b^*(\mathcal{A}_a-\tilde{\mathcal{A}}_a)\|. \qquad (C.13)\end{aligned}$$

We can write

$$\|(\mathcal{A}-\tilde{\mathcal{A}})\alpha^\lambda\|_0\approx 2\|\mathcal{A}_a^*(\mathcal{A}_a-\tilde{\mathcal{A}}_a)\alpha^\lambda\|_0$$

$$\approx 2 \left\| \sum_{i,j}^{N_d,N_s} g_i^*(\cdot)\overline{g_j(\cdot)} \int_\Omega \left(g_j(\acute{\mathbf{x}})\overline{e_i^*(\acute{\mathbf{x}})} + \overline{g_i^*(\acute{\mathbf{x}})}e_j(\acute{\mathbf{x}}) \right) \alpha^\lambda(\acute{\mathbf{x}}) d\acute{\mathbf{x}} \right\|_0$$

$$\leq 2 \max_{i,j} \|g_i^* g_j\|_0 \sum_{i,j}^{N_d,N_s} \int_\Omega \left| \left(g_j(\acute{\mathbf{x}})\overline{e_i^*(\acute{\mathbf{x}})} + \overline{g_i^*(\acute{\mathbf{x}})}e_j(\acute{\mathbf{x}}) \right) \alpha^\lambda(\acute{\mathbf{x}}) \right| d\acute{\mathbf{x}}. \tag{C.14}$$

An upper bound for the integral in (C.14) can be obtained as follows:

$$\int_\Omega \left| \left(g_j(\acute{\mathbf{x}})\overline{e_i^*(\acute{\mathbf{x}})} + \overline{g_i^*(\acute{\mathbf{x}})}e_j(\acute{\mathbf{x}}) \right) \alpha^\lambda(\acute{\mathbf{x}}) \right| d\acute{\mathbf{x}}$$
$$\leq \sum_{n=1}^{N_\Delta^{*i}} \|e_i^*\|_{0,ni} \|g_j \alpha^\lambda\|_{0,ni} + \sum_{m=1}^{N_\Delta^j} \|e_j\|_{0,mj} \|g_i^* \alpha^\lambda\|_{0,mj}. \tag{C.15}$$

Using (C.15) in (C.14),

$$\|(\mathcal{A} - \tilde{\mathcal{A}})\alpha^\lambda\|_0 \leq 2 \max_{i,j} \|g_i^* g_j\|_0$$
$$\times \left(\sum_{i=1}^{N_d} \sum_{n,j}^{N_\Delta^{*i},N_s} \|e_i^*\|_{0,ni} \|g_j \alpha^\lambda\|_{0,ni} + \sum_{j=1}^{N_s} \sum_{m,i}^{N_\Delta^j,N_d} \|e_j\|_{0,mj} \|g_i^* \alpha^\lambda\|_{0,mj} \right). \tag{C.16}$$

Similarly,

$$\|(\mathcal{A}_\mathcal{B} - \tilde{\mathcal{A}}_\mathcal{B})\beta^\lambda\|_0 \leq 2 \max_{i,j} \|g_i^* g_j\|_0$$
$$\times \left(\sum_{i=1}^{N_d} \sum_{n,j}^{N_\Delta^{*i},N_s} \|\nabla e_i^*\|_{0,ni} \left\| |\nabla g_j| \beta^\lambda \right\|_{0,ni} + \sum_{j=1}^{N_s} \sum_{m,i}^{N_\Delta^j,N_d} \|\nabla e_j\|_{0,mj} \left\| |\nabla g_i^*| \beta^\lambda \right\|_{0,mj} \right). \tag{C.17}$$

$$\|(\mathcal{B}_\mathcal{A} - \widetilde{\mathcal{B}_\mathcal{A}})\alpha^\lambda\|_0 \leq 2 \max_{i,j} \|\nabla g_i^* \cdot \nabla g_j\|_0$$
$$\times \left(\sum_{i=1}^{N_d} \sum_{n,j}^{N_\Delta^{*i},N_s} \|e_i^*\|_{0,ni} \left\| g_j \alpha^\lambda \right\|_{0,ni} + \sum_{j=1}^{N_s} \sum_{m,i}^{N_\Delta^j,N_d} \|e_j\|_{0,mj} \left\| g_i^* \alpha^\lambda \right\|_{0,mj} \right). \tag{C.18}$$

$$\|(\mathcal{B} - \tilde{\mathcal{B}})\beta^\lambda\|_0 \leq 2 \max_{i,j} \|\nabla g_i^* \cdot \nabla g_j\|_0$$
$$\times \left(\sum_{i=1}^{N_d} \sum_{n,j}^{N_\Delta^{*i},N_s} \|\nabla e_i^*\|_{0,ni} \left\| |\nabla g_j| \beta^\lambda \right\|_{0,ni} + \sum_{j=1}^{N_s} \sum_{m,i}^{N_\Delta^j,N_d} \|\nabla e_j\|_{0,mj} \left\| |\nabla g_i^*| \beta^\lambda \right\|_{0,mj} \right).$$

(C.19)

$\|\gamma_a - \tilde{\gamma}_a\|_0$ can be interpreted as follows:

$$\begin{aligned}\|\gamma_a - \tilde{\gamma}_a\|_0 &= \left\| \sum_{i,j}^{N_d,N_s} \left(g_i^*(\cdot)\overline{g_j(\cdot)} - G_i^*(\cdot)\overline{G_j(\cdot)} \right) \Gamma_{i,j} \right\|_0 \\ &\approx \left\| \sum_{i,j}^{N_d,N_s} \left(e_i^*(\cdot)\overline{g_j(\cdot)} + g_i^*(\cdot)\overline{e_j(\cdot)} \right) \Gamma_{i,j} \right\|_0 , \end{aligned} \tag{C.20}$$

where the error in $\Gamma_{i,j}$ due to discretization is neglected and the last approximation is derived similar to (C.10). Similarly,

$$\|\gamma_b - \tilde{\gamma}_b\|_0 \approx \left\| \sum_{i,j}^{N_d,N_s} \left(\nabla e_i^*(\cdot) \cdot \nabla\overline{g_j(\cdot)} + \nabla g_i^*(\cdot) \cdot \nabla\overline{e_j(\cdot)} \right) \Gamma_{i,j} \right\|_0 , \tag{C.21}$$

To compute an upper bound for $\|\tilde{\gamma}_a - \gamma_a\|_0$ using (C.20), we first write

$$\begin{aligned} &\left\| \sum_{i,j}^{N_d,N_s} \left(e_i^*(\cdot)\overline{g_j(\cdot)} + g_i^*(\cdot)\overline{e_j(\cdot)} \right) \Gamma_{i,j} \right\|_0 \\ &\qquad \le \max_{i,j} |\Gamma_{i,j}| \left\| \sum_{i,j}^{N_d,N_s} \left(e_i^*(\cdot)\overline{g_j(\cdot)} + g_i^*(\cdot)\overline{e_j(\cdot)} \right) \right\|_0 \\ &\le \max_{i,j} |\Gamma_{i,j}| \left(\sum_{i=1}^{N_d} \sum_{n,j}^{N_\Delta^{*i},N_s} \|e_i^*\|_{0,ni} \|g_j\|_{\infty,ni} + \sum_{i=j}^{N_s} \sum_{m,i}^{N_\Delta^{j},N_d} \|g_i^*\|_{\infty,mj} \|e_j\|_{0,mj} \right) . \end{aligned} \tag{C.22}$$

Noting (5.9),

$$\begin{aligned} \max_{i,j} |\Gamma_{i,j}| &\le \max_{i,j} \|g_i^* g_j\|_0 \|\alpha\|_0 + \max_{i,j} \|\nabla g_i^* \cdot \nabla g_j\|_0 \|\beta\|_0 , \\ &\le \max_{i,j} \|g_i^* g_j\|_1 \left(\|\alpha\|_0 + \|\beta\|_0 \right) . \end{aligned} \tag{C.23}$$

(C.23) leads to

$$\|\gamma_a - \tilde{\gamma}_a\|_0 \le \max_{i,j} \|g_i^* g_j\|_1 \left(\|\alpha\|_0 + \|\beta\|_0 \right)$$

$$\times \left(\sum_{i=1}^{N_d} \sum_{n,j}^{N_\Delta^{*i}, N_s} \|e_i^*\|_{0,ni} \|g_j\|_{\infty,ni} + \sum_{i=j}^{N_s} \sum_{m,i}^{N_\Delta^{j}, N_d} \|g_i^*\|_{\infty,mj} \|e_j\|_{0,mj} \right) . \tag{C.24}$$

Similarly,

$$\begin{aligned} \|\gamma_b - \tilde{\gamma}_b\|_0 \leq \max_{i,j} \|g_i^* g_j\|_1 \left(\|\alpha\|_0 + \|\beta\|_0 \right) \\ \times \left(\sum_{i=1}^{N_d} \sum_{n,j}^{N_\Delta^{*i}, N_s} \|\nabla e_i^*\|_{0,ni} \|\nabla g_j\|_{\infty,ni} + \sum_{i=j}^{N_s} \sum_{m,i}^{N_\Delta^{j}, N_d} \|\nabla g_i^*\|_{\infty,mj} \|\nabla e_j\|_{0,mj} \right) . \end{aligned} \tag{C.25}$$

Since,

$$\begin{aligned} \|\alpha^\lambda - \tilde{\alpha}^\lambda\| \leq \tilde{\chi}_{11} \left(\|(\tilde{\mathcal{A}} - \mathcal{A})\alpha\| + \|(\tilde{\mathcal{A}}_B - \mathcal{A}_B)\beta\| + \|\tilde{\gamma}_a - \gamma_a\| \right) \\ + \tilde{\chi}_{12} \left(\|(\tilde{\mathcal{B}}_A - \mathcal{B}_A)\alpha\| + \|(\tilde{\mathcal{B}} - \mathcal{B})\beta\| + \|\tilde{\gamma}_b - \gamma_b\| \right) , \end{aligned} \tag{C.26}$$

$$\begin{aligned} \|\beta^\lambda - \tilde{\beta}^\lambda\| \leq \tilde{\chi}_{21} \left(\|(\tilde{\mathcal{A}} - \mathcal{A})\alpha\| + \|(\tilde{\mathcal{A}}_B - \mathcal{A}_B)\beta\| + \|\tilde{\gamma}_a - \gamma_a\| \right) \\ + \tilde{\chi}_{22} \left(\|(\tilde{\mathcal{B}}_A - \mathcal{B}_A)\alpha\| + \|(\tilde{\mathcal{B}} - \mathcal{B})\beta\| + \|\tilde{\gamma}_b - \gamma_b\| \right) . \end{aligned} \tag{C.27}$$

Then,

$$\begin{aligned} \|\alpha^\lambda - \tilde{\alpha}^\lambda\|_0 \leq 2\tilde{\chi}_{11} \max_{i,j} \|g_i^* g_j\|_1 \\ \times \Bigg(\sum_{i=1}^{N_d} \sum_{n,j}^{N_\Delta^{*i}, N_s} \|e_i^*\|_{0,ni} \left(\|g_j \alpha^\lambda\|_{0,ni} + \frac{\|\alpha\|_0 + \|\beta\|_0}{2} \|g_j\|_{\infty,ni} \right) + \|\nabla e_i^*\|_{0,ni} \left\| |\nabla g_j| \beta^\lambda \right\|_{0,ni} \\ + \sum_{j=1}^{N_s} \sum_{m,i}^{N_\Delta^{j}, N_d} \|e_j\|_{0,mj} \left(\|g_i^* \alpha^\lambda\|_{0,mj} + \frac{\|\alpha\|_0 + \|\beta\|_0}{2} \|g_i^*\|_{\infty,mj} \right) + \|\nabla e_j\|_{0,mj} \left\| |\nabla g_i^*| \beta^\lambda \right\|_{0,mj} \Bigg) . \\ + 2\tilde{\chi}_{12} \max_{i,j} \|g_i^* g_j\|_1 \\ \times \Bigg(\sum_{i=1}^{N_d} \sum_{n,j}^{N_\Delta^{*i}, N_s} \|e_i^*\|_{0,ni} \|g_j \alpha^\lambda\|_{0,ni} + \|\nabla e_i^*\|_{0,ni} \left(\left\| |\nabla g_j| \beta^\lambda \right\|_{0,ni} + \frac{\|\alpha\|_0 + \|\beta\|_0}{2} \|\nabla g_j\|_{\infty,ni} \right) \\ + \sum_{j=1}^{N_s} \sum_{m,i}^{N_\Delta^{j}, N_d} \|e_j\|_{0,mj} \|g_i^* \alpha^\lambda\|_{0,mj} + \|\nabla e_j\|_{0,mj} \left(\left\| |\nabla g_i^*| \beta^\lambda \right\|_{0,mj} + \frac{\|\alpha\|_0 + \|\beta\|_0}{2} \|\nabla g_i\|_{\infty,mj} \right) \Bigg) , \end{aligned} \tag{C.28}$$

which can be organized as

$$
\begin{aligned}
&\|\alpha^\lambda - \tilde{\alpha}^\lambda\|_0 \leq 2 \max_{i,j} \|g_i^* g_j\|_1 \times \\
&\Bigg\{ \sum_{j=1}^{N_s} \sum_{m,i}^{N_\Delta^j, N_d} \Bigg[\|e_j\|_{0,mj} \Bigg((\tilde{\chi}_{11} + \tilde{\chi}_{12}) \|g_i^* \alpha^\lambda\|_{0,mj} + \tilde{\chi}_{12} \frac{\|\alpha\|_\infty + \|\beta\|_\infty}{2} \|g_i^*\|_{\infty,mj} \Bigg) \\
&\quad + \|\nabla e_j\|_{0,mj} \Bigg((\tilde{\chi}_{11} + \tilde{\chi}_{12}) \big\| |\nabla g_i^*| \beta^\lambda \big\|_{0,mj} + \tilde{\chi}_{12} \frac{\|\alpha\|_0 + \|\beta\|_0}{2} \|\nabla g_i^*\|_{\infty,mj} \Bigg) \Bigg] \\
&\quad + \sum_{i=1}^{N_d} \sum_{n,j}^{N_\Delta^{*i}, N_s} \Bigg[\|e_i^*\|_{0,ni} \Bigg((\tilde{\chi}_{11} + \tilde{\chi}_{12}) \|g_j \alpha^\lambda\|_{0,ni} + \tilde{\chi}_{12} \frac{\|\alpha\|_0 + \|\beta\|_0}{2} \|g_j\|_{\infty,ni} \Bigg) \\
&\quad + \|\nabla e_i^*\|_{0,ni} \Bigg((\tilde{\chi}_{11} + \tilde{\chi}_{12}) \big\| |\nabla g_j| \beta^\lambda \big\|_{0,ni} + \tilde{\chi}_{12} \frac{\|\alpha\|_0 + \|\beta\|_0}{2} \|\nabla g_j\|_{\infty,ni} \Bigg) \Bigg] \Bigg\}
\end{aligned}
\tag{C.29}
$$

Using the discretization error estimates (5.25)-(5.26) with the generic positive constant C, we can rewrite the upper bound as follows:

$$
\begin{aligned}
&\|\alpha^\lambda - \tilde{\alpha}^\lambda\|_0 \leq 2C \max_{i,j} \|g_i^* g_j\|_1 \times \\
&\Bigg\{ \sum_{j=1}^{N_s} \sum_{m,i}^{N_\Delta^j, N_d} \Bigg[h_{mj}^2 \|g_j\|_{2,m} \Bigg((\tilde{\chi}_{11} + \chi_{12}) \|g_i^* \alpha^\lambda\|_{0,mj} + \tilde{\chi}_{12} \frac{\|\alpha\|_0 + \|\beta\|_0}{2} \|g_i^*\|_{\infty,mj} \Bigg) \\
&\quad + h_{mj} \|g_j\|_{2,m} \Bigg((\tilde{\chi}_{11} + \tilde{\chi}_{12}) \big\| |\nabla g_i^*| \beta^\lambda \big\|_{0,mj} + \tilde{\chi}_{12} \frac{\|\alpha\|_0 + \|\beta\|_0}{2} \|\nabla g_i^*\|_{\infty,mj} \Bigg) \Bigg] \\
&\quad + \sum_{i=1}^{N_d} \sum_{n,j}^{N_\Delta^{*i}, N_s} \Bigg[h_{ni}^2 \|g_i^*\|_{2,ni} \Bigg((\tilde{\chi}_{11} + \tilde{\chi}_{12}) \|g_j \alpha^\lambda\|_{0,ni} + \tilde{\chi}_{12} \frac{\|\alpha\|_0 + \|\beta\|_0}{2} \|g_j\|_{\infty,ni} \Bigg) \\
&\quad + h_{ni} \|g_i^*\|_{2,ni} \Bigg((\tilde{\chi}_{11} + \tilde{\chi}_{12}) \big\| |\nabla g_j| \beta^\lambda \big\|_{0,ni} + \tilde{\chi}_{12} \frac{\|\alpha\|_0 + \|\beta\|_0}{2} \|\nabla g_j\|_{\infty,ni} \Bigg) \Bigg] \Bigg\}.
\end{aligned}
$$

Using a similar approach, the bound for $\|\beta^\lambda - \tilde{\beta}^\lambda\|_0$ can be found to be

$$
\begin{aligned}
&\|\beta^\lambda - \tilde{\beta}^\lambda\|_0 \leq 2C \max_{i,j} \|g_i^* g_j\|_1 \times \\
&\Bigg\{ \sum_{j=1}^{N_s} \sum_{m,i}^{N_\Delta^j, N_d} \Bigg[h_{mj}^2 \|g_j\|_{2,m} \Bigg((\tilde{\chi}_{21} + \tilde{\chi}_{22}) \|g_i^* \alpha^\lambda\|_{0,mj} + \tilde{\chi}_{22} \frac{\|\alpha\|_0 + \|\beta\|_0}{2} \|g_i^*\|_{\infty,mj} \Bigg)
\end{aligned}
$$

$$
\begin{aligned}
&+h_{mj}\|g_j\|_{2,m}\left((\tilde{\chi}_{21}+\tilde{\chi}_{22})\left\||\nabla g_i^*|\beta^\lambda\right\|_{0,mj}+\tilde{\chi}_{22}\frac{\|\alpha\|_0+\|\beta\|_0}{2}\|\nabla g_i^*\|_{\infty,mj}\right)\Bigg]\\
+\sum_{i=1}^{N_d}\sum_{n,j}^{N_\Delta^{*i},N_s}&\left[h_{ni}^2\|g_i^*\|_{2,ni}\left((\tilde{\chi}_{21}+\tilde{\chi}_{22})\|g_j\alpha^\lambda\|_{0,ni}+\tilde{\chi}_{22}\frac{\|\alpha\|_0+\|\beta\|_0}{2}\|g_j\|_{\infty,ni}\right)\right.\\
&\left.+h_{ni}\|g_i^*\|_{2,ni}\left((\tilde{\chi}_{21}+\tilde{\chi}_{22})\left\||\nabla g_j|\beta^\lambda\right\|_{0,ni}+\tilde{\chi}_{22}\frac{\|\alpha\|_0+\|\beta\|_0}{2}\|\nabla g_j\|_{\infty,ni}\right)\right]\Bigg\}.
\end{aligned}
$$

Letting $a(j,m), b(j,m), c(j,m)$ and $d(j,m)$ as follows yields the first part of the theorem:

$$
\begin{aligned}
a(j,m) &:= \sum_{i=1}^{N_d}\|g_i^*\alpha^\lambda\|_{0,mj}\|g_j\|_{2,mj}\\
b(j,m) &:= \frac{\|\alpha\|_0+\|\beta\|_0}{2}\sum_{i=1}^{N_d}\|g_i^*\|_{\infty,mj}\|g_j\|_{2,mj},\\
c(j,m) &:= \sum_{i=1}^{N_d}\left\||\nabla g_i^*|\beta^\lambda\right\|_{0,mj}\|g_j\|_{2,mj}\\
d(j,m) &:= \frac{\|\alpha\|_0+\|\beta\|_0}{2}\sum_{j=1}^{N_s}\|\nabla g_i^*\|_{\infty,mj}\|g_j\|_{2,mj},
\end{aligned}
$$

Defining $a^*(j,m), b^*(j,m), c^*(j,m)$ and $d^*(j,m)$ in a similar way concludes the theorem.

C.2 Proof of theorem 2

Clearly,

$$
\begin{aligned}
\tilde{\sigma}^\lambda-\tilde{\sigma}^\lambda_{n,m} &= \left[\mathcal{I}-(\mathcal{P}^{n,m}\tilde{\mathcal{T}})^{-1}\mathcal{P}^{n,m}\tilde{\mathcal{T}}\right]\tilde{\sigma}^\lambda\\
&= \left[\mathcal{I}-(\mathcal{P}^{n,m}\tilde{\mathcal{T}})^{-1}\mathcal{P}^{n,m}\tilde{\mathcal{T}}\right]\left(\tilde{\sigma}^\lambda-\psi\right),
\end{aligned}
$$

where $\psi := [\psi_a\ \psi_b]^T \in X^a \times X^b$. Note that $\left[\mathcal{I}-(\mathcal{P}^{n,m}\tilde{\mathcal{T}})^{-1}\mathcal{P}^{n,m}\tilde{\mathcal{T}}\right]\psi = 0$.

Without loss of generality, we can express $(\mathcal{P}^{n,m}\tilde{\mathcal{T}})^{-1}\mathcal{P}^{n,m}$ as follows:

$$
(\mathcal{P}^{n,m}\tilde{\mathcal{T}})^{-1}\mathcal{P}^{n,m} = \begin{pmatrix}\Pi_{11} & \Pi_{12}\\ \Pi_{21} & \Pi_{22}\end{pmatrix}, \tag{C.30}
$$

where Π_{ij}, for $i = 1, 2$, $j = 1, 2$ are defined on appropriate subspaces. The boundedness of $(\mathcal{P}^{n,m}\tilde{\mathcal{T}})^{-1}$ implies that Π_{ij} is bounded by

$$\|\Pi_{ij}\| \le \pi_{ij} \tag{C.31}$$

for some positive π_{ij}, $i = 1, 2$, $j = 1, 2$. Note that $\|(\mathcal{P}^{n,m}\tilde{\mathcal{T}})^{-1}\mathcal{P}^{n,m}\| \le \|(\mathcal{P}^{n,m}\tilde{\mathcal{T}})^{-1}\|$, since $\mathcal{P}^{n,m}$ is the orthogonal projection operator and $\|\mathcal{P}^{n,m}\| = 1$ [57].

Let $e_a := \tilde{\alpha}^\lambda - \psi_a$ and $e_b := \tilde{\beta}^\lambda - \psi_b$. We write

$$\tilde{\alpha}^\lambda - \tilde{\alpha}_n^\lambda = (e_a - \lambda_a \Pi_{11} e_a - \Pi_{12}\lambda_b e_b) - \Pi_{11}(\tilde{\mathcal{A}} e_a + \tilde{\mathcal{A}}_B e_b) - \Pi_{12}(\tilde{\mathcal{B}}_A e_a + \tilde{\mathcal{B}} e_b) \tag{C.32}$$

Then,

$$\begin{aligned}\|\tilde{\alpha}^\lambda - \tilde{\alpha}_n^\lambda\|_0 \le{} & \|(\mathcal{I} - \lambda_a \Pi_{11})e_a\|_0 + \lambda_b \|\Pi_{12} e_b\|_0 \\ & \|\Pi_{11}\tilde{\mathcal{A}} e_a + \Pi_{12}\tilde{\mathcal{B}}_A e_a\|_0 + \|\Pi_{11}\tilde{\mathcal{A}}_B e_b + \Pi_{12}\tilde{\mathcal{B}} e_b\|_0.\end{aligned}$$

We have

$$\begin{aligned}\|\tilde{\alpha}^\lambda - \tilde{\alpha}_n^\lambda\|_0 \le{} & \|\mathcal{I} - \lambda_a \Pi_{11}\|_{L^2(\Omega)\to L^2(\Omega)} \|e_a\|_0 + \pi_{11}\|\tilde{\mathcal{A}} e_a\|_0 + \pi_{12}\|\tilde{\mathcal{B}}_A e_a\|_0 \\ & + \pi_{11}\|\tilde{\mathcal{A}}_B e_b\|_0 + \pi_{12}\|\tilde{\mathcal{B}} e_b\|_0.\end{aligned}$$

Assuming further that $\tilde{\alpha}^\lambda, \tilde{\beta}^\lambda$ are bounded on $\Omega \bigcup \partial\Omega$, and recalling the interpolation error estimates [20]

$$\begin{aligned}\|e_\alpha\|_0 &\le C\|\tilde{\alpha}^\lambda\|_1 h_{ta}, \\ \|e_\beta\|_0 &\le C\|\tilde{\beta}^\lambda\|_1 h_{ub},\end{aligned}$$

for some generic $C > 0$ and using the results of theorem 1 [39], we can write:

$$\begin{aligned}\|\tilde{\alpha}^\lambda - \tilde{\alpha}_n^\lambda\|_0 \le{} & C(1 + \lambda_a \pi_{11}) \sum_{t=1}^{N_a^\Delta} \|\tilde{\alpha}^\lambda\|_{1,ta} h_{ta} \\ & + C\lambda_b \pi_{12} \sum_{u=1}^{N_b^\Delta} \|\tilde{\beta}^\lambda\|_{1,ub} h_{ub}\end{aligned}$$

$$+C\pi_{11}\max_{i,j}\|G_i^*G_j\|_0\sum_{t=1}^{N_a^\Delta}\sum_{i,j}^{N_d,N_s}\|G_i^*G_j\|_{0,ta}\|\tilde{\alpha}^\lambda\|_{1,ta}h_{ta}$$

$$+C\pi_{12}\max_{i,j}\|\nabla G_i^*\cdot\nabla G_j\|_0\sum_{t=1}^{N_a^\Delta}\sum_{i,j}^{N_d,N_s}\|G_i^*G_j\|_{0,ta}\|\tilde{\alpha}^\lambda\|_{1,ta}h_{ta}$$

$$+C\pi_{11}\max_{i,j}\|G_i^*G_j\|_0\sum_{u=1}^{N_b^\Delta}\sum_{i,j}^{N_d,N_s}\|\nabla G_i^*\cdot\nabla G_j\|_{0,ub}\|\tilde{\beta}^\lambda\|_{1,ub}h_{ub}$$

$$+C\pi_{12}\max_{i,j}\|\nabla G_i^*\cdot\nabla G_j\|_0\sum_{u=1}^{N_b^\Delta}\sum_{i,j}^{N_d,N_s}\|\nabla G_i^*\cdot\nabla G_j\|_{0,ub}\|\tilde{\beta}^\lambda\|_{1,ub}h_{ub}$$

which can be put into a simpler form as follows:

$$\begin{aligned}\|\tilde{\alpha}^\lambda-\tilde{\alpha}_n^\lambda\|_0 \le{}& C(1+\lambda_a\pi_{11})\sum_{t=1}^{N_a^\Delta}\|\tilde{\alpha}^\lambda\|_{1,ta}h_{ta}+C\lambda_b\pi_{12}\sum_{u=1}^{N_b^\Delta}\|\tilde{\beta}^\lambda\|_{1,ub}h_{ub}\\ &+C\left(\pi_{11}\max_{i,j}\|G_i^*G_j\|_0+\pi_{12}\max_{i,j}\|\nabla G_i^*\cdot\nabla G_j\|_0\right)\\ &\times\sum_{t=1}^{N_a^\Delta}\sum_{i,j}^{N_d,N_s}\|G_i^*G_j\|_{0,ta}\|\tilde{\alpha}^\lambda\|_{1,ta}h_{ta}\\ &+C\left(\pi_{11}\max_{i,j}\|G_i^*G_j\|_0+\pi_{12}\max_{i,j}\|\nabla G_i^*\cdot\nabla G_j\|_0\right)\\ &\times\sum_{u=1}^{N_b^\Delta}\sum_{i,j}^{N_d,N_s}\|\nabla G_i^*\cdot\nabla G_j\|_{0,ub}\|\tilde{\beta}^\lambda\|_{1,ub}h_{ub}\end{aligned}$$

Since

$$\max_{i,j}\|G_i^*G_j\|_0+\max_{i,j}\|\nabla G_i^*\cdot\nabla G_j\|_0\le\max_{i,j}\|G_i^*G_j\|_1, \tag{C.33}$$

we write

$$\begin{aligned}\|\tilde{\alpha}^\lambda-\tilde{\alpha}_n^\lambda\|_0 \le{}& C(1+\lambda_a\pi_{11})\sum_{t=1}^{N_a^\Delta}\|\tilde{\alpha}^\lambda\|_{1,ta}h_{ta}+C\lambda_b\pi_{12}\sum_{u=1}^{N_b^\Delta}\|\tilde{\beta}^\lambda\|_{1,ub}h_{ub}\\ &+C(\pi_{11}+\pi_{12})\max_{i,j}\|G_i^*G_j\|_1\sum_{t=1}^{N_a^\Delta}\sum_{i,j}^{N_d,N_s}\|G_i^*G_j\|_{0,ta}\|\tilde{\alpha}^\lambda\|_{1,ta}h_{ta}\end{aligned}$$

$$
\begin{aligned}
&+C(\pi_{11}+\pi_{12})\max_{i,j}\|G_i^*G_j\|_1\sum_{u=1}^{N_b^\Delta}\sum_{i,j}^{N_d,N_s}\|\nabla G_i^*\cdot\nabla G_j\|_{0,ub}\|\tilde{\beta}^\lambda\|_{1,ub}h_{ub}\\
&\leq C(1+\lambda_a\pi_{11})\sum_{t=1}^{N_a^\Delta}\|\tilde{\alpha}^\lambda\|_{1,ta}h_{ta}+C\lambda_b\pi_{12}\sum_{u=1}^{N_b^\Delta}\|\tilde{\beta}^\lambda\|_{1,ub}h_{ub}+C(\pi_{11}+\pi_{12})\max_{i,j}\|G_i^*G_j\|_1\\
&\quad\times\left(\sum_{t=1}^{N_a^\Delta}\sum_{i,j}^{N_d,N_s}\|G_i^*G_j\|_{0,ta}\|\tilde{\alpha}^\lambda\|_{1,ta}h_{ta}+\sum_{u=1}^{N_b^\Delta}\sum_{i,j}^{N_d,N_s}\|\nabla G_i^*\cdot\nabla G_j\|_{0,ub}\|\tilde{\beta}^\lambda\|_{1,ub}h_{ub}\right).
\end{aligned}
$$

A similar bound can be obtained for $\|\tilde{\beta}^\lambda-\tilde{\beta}_m^\lambda\|_0$:

$$
\begin{aligned}
\|\tilde{\beta}^\lambda-\tilde{\beta}_m^\lambda\|_0 &\leq C(1+\lambda_b\pi_{22})\sum_{u=1}^{N_b^\Delta}\|\tilde{\beta}^\lambda\|_{1,ub}h_{ub}+C\lambda_a\pi_{21}\sum_{u=1}^{N_a^\Delta}\|\tilde{\alpha}^\lambda\|_{1,ta}h_{ta}\\
&\quad+C(\pi_{21}+\pi_{22})\max_{i,j}\|G_i^*G_j\|_1\\
&\times\left(\sum_{t=1}^{N_a^\Delta}\sum_{i,j}^{N_d,N_s}\|G_i^*G_j\|_{0,ta}\|\tilde{\alpha}^\lambda\|_{1,ta}h_{ta}+\sum_{u=1}^{N_b^\Delta}\sum_{i,j}^{N_d,N_s}\|\nabla G_i^*\cdot\nabla G_j\|_{0,ub}\|\tilde{\beta}^\lambda\|_{1,ub}h_{ub}\right).
\end{aligned}
$$

C.3 Proof of theorem 3

The conditions (5.45) and (5.46) together with (5.47)-(5.48) imply that

$$
\begin{aligned}
\left[(\tilde{\chi}_{11}+\tilde{\chi}_{12})a(j,m)+\tilde{\chi}_{12}b(j,m)\right]h_{mj}^2+\left[(\tilde{\chi}_{11}+\tilde{\chi}_{12})c(j,m)+\tilde{\chi}_{12}d(j,m)\right]h_{mj} &\leq \varepsilon,\\
\left[(\tilde{\chi}_{11}+\tilde{\chi}_{12})a^*(i,n)+\tilde{\chi}_{12}b^*(i,n)\right]h_{ni}^2+\left[(\tilde{\chi}_{11}+\tilde{\chi}_{12})c^*(i,n)+\tilde{\chi}_{12}d^*(i,n)\right]h_{ni} &\leq \varepsilon,
\end{aligned}
$$

and

$$
\begin{aligned}
\left[(\tilde{\chi}_{21}+\tilde{\chi}_{22})a(j,m)+\tilde{\chi}_{22}b(j,m)\right]h_{mj}^2+\left[(\tilde{\chi}_{21}+\tilde{\chi}_{22})c(j,m)+\tilde{\chi}_{22}d(j,m)\right]h_{mj} &\leq \varepsilon,\\
\left[(\tilde{\chi}_{21}+\tilde{\chi}_{22})a^*(i,n)+\tilde{\chi}_{22}b^*(i,n)\right]h_{ni}^2+\left[(\tilde{\chi}_{21}+\tilde{\chi}_{22})c^*(i,n)+\tilde{\chi}_{22}d^*(i,n)\right]h_{ni} &\leq \varepsilon,
\end{aligned}
$$

since $a(j,m)$, $b(j,m)$, $c(i,n)$, $d(i,n)$ and $a^*(j,m)$, $b^*(j,m)$, $c^*(i,n)$, $d^*(i,n)$ are positive. Using this argument in theorem 1 results in the bound given in (5.44).

APPENDIX D
Appendix

D.1 Finite Element Discretization of the Forward Problem

In this section, we formulate the finite element problem for the diffusion equation in frequency domain on a bounded domain $\Omega \subset \mathbb{R}^2$. We use triangular finite elements with piece-wise linear Lagrange basis functions for the discretization.

We turn back to the boundary value problem in equations (**??**) and (**??**) (dropping the subscript j from g):

$$-\nabla \cdot D(\mathbf{x})\nabla g + \left(\mu_a(\mathbf{x}) + \frac{i\omega}{c}\right) g = Q(\mathbf{x}) \quad (\mathbf{x}) \in \Omega, \tag{D.1}$$

$$g + 2aD(\mathbf{x})\frac{\partial g}{\partial n} = 0 \quad (\mathbf{x}) \in \partial\Omega, \tag{D.2}$$

where $\Omega \subset \mathbb{R}^2$ is a bounded domain with Lipschitz boundary $\partial\Omega$ and outward unit normal $\hat{n}$.

Equivalently we can express the problem in equations (D.1)-(D.2) as follows:

$$-(D(\mathbf{x})g_x)_x - (D(\mathbf{x})g_y)_y + \left(\mu_a(\mathbf{x}) + \frac{j\omega}{c}\right) g = Q(\mathbf{x}), \qquad (\mathbf{x}) \in \Omega \tag{D.3}$$

$$g(\mathbf{x}) + 2aD(\mathbf{x})\hat{n} \cdot \nabla g(\mathbf{x}) = 0 \qquad (\mathbf{x}) \in \partial\Omega \tag{D.4}$$

We first formulate the variational problem by multiplying equation (D.3) by a test function $\overline{\psi} \in H^1$ and integrating over Ω to obtain

$$\int_\Omega \overline{\psi}\left[(Dg_x)_x + (Dg_y)_y + \left(\mu_a + \frac{j\omega}{c}\right) g - Q\right] d\mathbf{x} = 0. \tag{D.5}$$

Applying divergence theorem, we get

$$\begin{aligned} \int_\Omega \left[\nabla\overline{\psi} \cdot D\nabla g + \overline{\psi}\left((\mu_a + \frac{j\omega}{c})g - Q\right)\right] d\mathbf{x} - \int_{\partial\Omega} \overline{\psi} D \frac{\partial g}{\partial n} dl &= 0, \\ \int_\Omega \left[\nabla\overline{\psi} \cdot D\nabla g + \overline{\psi}\left((\mu_a + \frac{j\omega}{c})g - Q\right)\right] d\mathbf{x} + \int_{\partial\Omega} \overline{\psi} \frac{1}{2a} g dl &= 0, \end{aligned} \tag{D.6}$$

where the integrand Dg_n in the boundary integral is replaced by $g/(2a)$ using the boundary condition in (D.2).

Equivalently we can express (D.6) as follows:

$$A(\psi, g) + \left\langle \psi, \frac{1}{2a} g \right\rangle = (\psi, Q) \tag{D.7}$$

where

$$\begin{aligned} A(\psi, g) &:= \int_\Omega \left[D(\overline{\psi}_x g_x + \overline{\psi}_y g_y) + \left(\mu_a + \frac{j\omega}{c} \right) \overline{\psi} g \right] d\mathbf{x}, && \text{(D.8)} \\ (\psi, Q) &:= \int_\Omega \overline{\psi} Q d\mathbf{x} && \text{(D.9)} \\ \left\langle \psi, \frac{1}{2a} g \right\rangle &:= \frac{1}{2a} \int_{\partial\Omega} \overline{\psi} g dl. && \text{(D.10)} \end{aligned}$$

Similarly, the variational problem can be formulated for the adjoint problem in (**??**) and (**??**) (dropping the subscript i) as follows :

$$A(\psi, g^*) + \left\langle \psi, \frac{1}{2a} g^* \right\rangle =< \psi, Q^* > \tag{D.11}$$

where Q^* is the adjoint source located at the detector position and ω is replaced by $-\omega$ in $A(\psi, g^*)$.

D.1.1 Existence and Uniqueness of the Solution to the Variational Problem

In this section, we prove that a solution to (D.7) (and to (D.11)) exists and it is unique. Furthermore, we show that the problem (D.7) is well-posed in the sense that the solution g is bounded.

We first write the following sesquilinear form [47]:

$$b(\psi, g) = A(\psi, g) + \left\langle \psi, \frac{1}{2a} g \right\rangle, \tag{D.12}$$

and consider

$$b(\psi, g) = (\psi, Q). \tag{D.13}$$

We start with the following proposition:

Proposition 1 For bounded $D \leq D_{max}$ and $\mu_a \leq \mu_{max}$, the sesquilinear form $b(\psi, g)$ is continuous, that is $|b(\psi, g)| \leq M\|\psi\|_1\|g\|_1$, for some $M > 0$.

Proof.

$$\begin{aligned}
|A(\psi, g)| &= \left|(\nabla\psi, D\nabla g) + \left(\psi, (\mu_a + \frac{j\omega}{c})g\right)\right| \\
&\leq \left|\int_\Omega \nabla\psi \cdot D\nabla g\right| + \left|\int_\Omega \psi\left(\mu_a + \frac{j\omega}{c}\right)g\right| \\
&\leq D_{max}\|\nabla\psi\|_0\|\nabla g\|_0 + \|\psi\|_0\|g\|_0|\mu_a + \frac{j\omega}{c}|_{max} \\
&\leq \|\psi\|_1\|g\|_1 D_{max} + \|\psi\|_1\|g\|_1\left|\mu_a + \frac{j\omega}{c}\right|_{max} \\
&= \|\psi\|_1\|g\|_1\left(D_{max} + \left|\mu_a + \frac{j\omega}{c}\right|_{max}\right) \\
&= M_1\|\psi\|_1\|g\|_1, \qquad \text{(D.14)}
\end{aligned}$$

where $M_1 = \left(D_{max} + \left|\mu_a + \frac{j\omega}{c}\right|_{max}\right)$. Note that

$$\left|\left\langle\psi, \frac{1}{2a}g\right\rangle\right| \leq \frac{1}{2a}\|\psi\|_{0,(\partial\Omega)}\|g\|_{0(\partial\Omega)}.$$

By [20]

$$\begin{aligned}
\frac{1}{2a}\|\psi\|_{0(\partial\Omega)}\|g\|_{0(\partial\Omega)} &\leq \frac{1}{2a}C_1\|\psi\|_0^{1/2}\|\psi\|_1^{1/2}C_2\|g\|_0^{1/2}\|g\|_1^{1/2} \\
&\leq \frac{1}{2a}M_2\|\psi\|_1\|g\|_1, \qquad \text{(D.15)}
\end{aligned}$$

where $M_2 = C_1C_2$, for some $C_1, C_2 > 0$. Thus, $|b(\psi, g)| \leq M\|\psi\|_1\|g\|_1$, where $M = M_1 + M_2/2a$. □

Proposition 2 For some bounded $D_{min} \leq D$ and bounded $\mu_{min} \leq \mu_a$, The sesquilinear form $b(\psi, g)$ is coercive (positive definite), that is $|b(\psi, \psi)| \geq \beta\|\psi\|_1^2$, for some $\beta > 0$.

Proof.

$$\begin{aligned}
|b(\psi,\psi)| &= \left|(\nabla\psi, D\nabla\psi) + \left(\psi, (\mu_a + \frac{j\omega}{c})\psi\right) + \left\langle \psi, \frac{1}{2a}\psi \right\rangle\right| \\
&= \left|\int_\Omega \nabla\psi \cdot D\nabla\psi + \int_\Omega \psi \left(\mu_a + \frac{j\omega}{c}\right)\psi + \frac{1}{2a}\int_{\partial\Omega} |\psi|^2\right| \\
&\geq D_{min}\int_\Omega |\nabla\psi|^2 + \left|(\mu_a + \frac{j\omega}{c})\right|_{min} \int_\Omega |\psi|^2 \\
&\geq \min\left(D_{min}, \left|\left(\mu_{min} + \frac{j\omega}{c}\right)\right|\right)\|\psi\|_1^2, \qquad \text{(D.16)}
\end{aligned}$$

where $\beta = \min(D_{min}, |(\mu_{min} + j\omega/c)|)$. □

By Lax-Milgram lemma, propositions 1 and 2 ensure that the variational problem (D.12) has a unique solution. Furthermore, we establish the following stability condition for the solution g of the frequency-domain diffusion equation:

Corollary The solution g to the variational formulation in equation (D.12) is stable in the sense that

$$\|g\|_1 \leq \frac{M}{\beta}, \qquad \text{(D.17)}$$

for some $C > 0$, which implies that g belongs to H^1.

Proof. The ellipticity condition in Proposition 2 implies that the solution g is bounded by the data f, measured in the norm of the dual space of H^1, which is H^{-1}. The bound on the norm of the solution is given by [47]

$$\|g\|_1 \leq \frac{1}{\beta}\|Q\|_{H^{-1}}, \qquad \text{(D.18)}$$

where the dual norm $\|Q\|_{H^{-1}}$ is given by

$$\|Q\|_{H^{-1}} := \sup_{0 \neq v \in H^1} \frac{|Q(v)|}{\|v\|_1}, \qquad \text{(D.19)}$$

where $Q(v) = (Q, v)$. At this point, we use the Gaussian approximation for the source Q (see Chapter 2), which implies that $Q(v) = (Q, v)$ is a bounded linear functional in H^{-1}. Thus, the dual norm of $Q(v)$ is given by

$$|Q(v)\|_{H^{-1}} := \sup_{0\neq v\in H^1} \frac{|Q(v)|}{\|v\|_1} \leq \sup_{0\neq v\in H^1} \frac{\|Q\|_0\|v\|_0}{\|v\|_1} \leq \sup_{0\neq v\in H^1} \frac{\|Q\|_0\|v\|_1}{\|v\|_1} = \|Q\|_0. \tag{D.20}$$

□

D.1.2 Finite Element Discretization

Consider the integrals in equations (D.8)-(D.9) and (D.10). In order to compute these integrals on a set of finite elements, we write the elemental integrals by restricting (D.8)-(D.9) and (D.10) on the finite elements as follows:

$$A(\psi, g) = \sum_{e=1}^{N_\Delta} A(g, \psi)_e, \tag{D.21}$$

$$(g, Q) = \sum_{e=1}^{N_\Delta} (\psi, Q)_e, \tag{D.22}$$

$$\left\langle \psi, \frac{1}{2a} g \right\rangle = \sum_{e=1}^{N_\Delta} \left\langle \psi, \frac{1}{2a} g \right\rangle_e . \tag{D.23}$$

where N_Δ is the number of finite elements.

These integrals may be simple or complex depending on μ_a, D, Q and the mesh. Pursuing a general approach, we can consider a procedure based on transforming these integrals on element e to a canonical element Ω_0 and evaluate them on the canonical element. Thus, we transform an arbitrary element in the physical (x, y)-plane to the canonical element having a simpler geometry in a computational (ξ, η)-plane. The details of this approach can be found elsewhere [9] Transformation leads to the following integrals:

$$\begin{aligned} A_e(\psi, g) &= \int_{\Omega_0} \big[D_0(\overline{\psi_{0_\xi}} \xi_x + \overline{\psi_{0_\eta}} \eta_x)(g_{0_\xi} \xi_x + g_{0_\eta \eta_x}) \\ &+ D_0(\overline{\psi_{0_\xi}} \xi_y + \overline{\psi_{0_\eta}} \eta_y)(g_{0_\xi} \xi_y + g_{0_\eta} \eta_y)(\mu_{a0} + \frac{j\omega}{c})\overline{\psi_0} g_0 \Big] \det(J_e) d\xi d\eta, \end{aligned} \tag{D.24}$$

$$(\psi, f)_e = \int_{\Omega_0} \overline{\psi_0(\xi, \eta)} f(x(\xi, \eta), y(\xi, \eta)) \det(J_e) d\xi d\eta, \tag{D.25}$$

$$< \psi, \frac{1}{2a} g >_e \; = \; \int_{\partial\Omega_0} \overline{\psi_0} \frac{1}{2a} g_0 dl, \tag{D.26}$$

where D_0, μ_{a0} are the diffusion and absorption coefficients transformed onto the canonical element, and J_e is the Jacobian of the transformations $x = x(\xi, \eta)$ and $y = y(\xi, \eta)$ and is given by

$$J_e = \begin{pmatrix} x_\xi & x_\eta \\ y_\xi & y_\eta \end{pmatrix}, \tag{D.27}$$

where the subscripts denote the variable with respect to which the derivative is taken.

D.1.3 Generation of Element Matrices and Vectors

In order to obtain the finite-element (Galerkin) formulation of (D.12), we use the elemental integrals (D.24)-(D.25) and (D.26), select a finite dimensional subspace $S^N \subset H^1$, and replace g and ψ with their finite-dimensional counterparts G and Ψ for $\mathbf{x} = (x, y)$ as follows:

$$G(\mathbf{x}) \; = \; \sum_{j=1}^{N} c_j \varphi_j(\mathbf{x}) \tag{D.28}$$

$$\; = \; \sum_{j=1}^{N} c_j \bigcup_{e=1}^{N_\Delta^j} N_{j,e}(\mathbf{x}) \tag{D.29}$$

$$\Psi(\mathbf{x}) \; = \; \sum_{j=1}^{N} d_j^* \bigcup_{e=1}^{N_\Delta^j} N_{j,e}(\mathbf{x}) \tag{D.30}$$

where $\varphi_j(\mathbf{x})$ denotes the j^{th} basis function associated with the j^{th} node and N is the number of nodes. c_j and d_j are the coefficients of each basis function in the finite sum for G and Ψ, respectively. $N_{j,e}$ is the elemental shape function [20] associated with the j^{th} node and N_Δ^j is the number of elements that include the j^{th} node. Note that we use piece-wise linear Lagrange polynomials as the basis functions and thus the Lagrange shape functions corresponding to these basis polynomials.

Transformation to the canonical (ξ, η) plane yields the following representations for the finite element approximations $G_0(\xi, \eta)$ and $\Psi_0(\xi, \eta)$ of $g(\mathbf{x})$ and $\psi(\mathbf{x})$, respectively on the canonical element.

$$\begin{aligned} G_0(\xi,\eta) &= \mathbf{c}_e^T\mathbf{N}(\xi,\eta) && \text{(D.31)} \\ \Psi_0(\xi,\eta) &= \mathbf{d}_e^T\mathbf{N}(\xi,\eta) && \text{(D.32)} \end{aligned}$$

where the vectors $\mathbf{c}_e$ and $\mathbf{d}_e$contain the elemental parameters and $\mathbf{N}(\xi,\eta)$ is a vector that contains the elemental shape functions:

$$\begin{aligned} \mathbf{c}_e^T &= \left[\begin{array}{cccc} c_{e,1} & c_{e,2} & \cdots & c_{e,N_e} \end{array}\right], && \text{(D.33)} \\ \mathbf{d}_e^T &= \left[\begin{array}{cccc} d_{e,1} & d_{e,2} & \cdots & d_{e,N_e} \end{array}\right], && \text{(D.34)} \\ \mathbf{N}_e^T &= \left[\begin{array}{cccc} N_1(\xi,\eta) & N_2(\xi,\eta) & \cdots & N_{N_e}(\xi,\eta) \end{array}\right], && \text{(D.35)} \end{aligned}$$

where N_e is the number of nodes in each element. N_e depends on the order of the basis selected for the finite element solution of the problem. Using the equivalent forms of $G_0(\xi,\eta)$ and $\Psi_0(\xi,\eta)$ given in equations (D.31) and (D.32) in the elemental integrals given in equations (D.24), (D.25) and (D.26), we arrive at

$$\begin{aligned} A(\Psi,G)_e &= \mathbf{d}_e^T(\mathbf{K}_e+\mathbf{M}_e)\mathbf{c}_e \\ (\Psi,Q)_e &= \mathbf{d}_e^T\mathbf{Q}_e \\ \left\langle \Psi,\frac{1}{2a}G\right\rangle_e &= \mathbf{d}_e^T\mathbf{P}_{e\partial}\mathbf{c}_e \end{aligned}$$

where

$$\begin{aligned} \mathbf{K}_e &= \int_{\Omega_0}\left[g_{1e}\mathbf{N}_\xi\mathbf{N}_\xi^T + g_{2e}(\mathbf{N}_\xi\mathbf{N}_\eta^T+\mathbf{N}_\eta\mathbf{N}_\xi^T) + g_{3e}\mathbf{N}_\eta\mathbf{N}_\eta^T\right]\det(J_e)d\xi d\eta \\ \mathbf{M}_e &= -\int_{\Omega_0}\left(\mu_{a0}+\frac{j\omega}{c}\right)\mathbf{N}\mathbf{N}^T\det(J_e)d\xi d\eta \\ \mathbf{Q}_e &= \int_{\Omega_0}\mathbf{N}Q\det(J_e)d\xi d\eta, \end{aligned}$$

with the following identities

$$g_{1e} = D_0[\xi_x^2+\xi_y^2]$$

$$
\begin{aligned}
g_{2e} &= D_0[\xi_x \eta_x + \xi_y \eta_y] \\
g_{3e} &= D_0[\eta_x^2 + \eta_y^2]
\end{aligned}
$$

and $D_0 = D(x(\xi,\eta), y(\xi,\eta))$ and $\mu_{a0} = \mu_a(x(\xi,\eta), y(\xi,\eta))$.

The term $\mathbf{P}_{e\partial}$ is computed according to where the boundary edge is transformed to. Assuming that the boundary edge of the physical element is transformed onto the edge of the canonical triangle with $\eta = 0$ via the transformation $x = x(\xi,\eta)$, $y = y(\xi,\eta)$, $\mathbf{P}_{e\partial}$ is given by

$$
\mathbf{P}_{e\partial} = \frac{1}{2a} \int_0^1 \mathbf{N}(\xi,0)\mathbf{N}^T(\xi,0) \sqrt{\left(\frac{dx(\xi,0)}{d\xi}\right)^2 + \left(\frac{dy(\xi,0)}{d\xi}\right)^2} d\xi. \tag{D.36}
$$

A similar approach can be followed for the case when the boundary edge is transformed onto $\xi = 0$. In case the boundary edge is transformed to the hypotenuse of the canonical element, the line integral can be evaluated as follows

$$
\mathbf{P}_{e\partial} = \frac{1}{2a} \int_0^1 \mathbf{N}(\xi,1-\xi)\mathbf{N}^T(\xi,1-\xi) \sqrt{\left(\frac{dx(\xi,1-\xi)}{d\xi}\right)^2 + \left(\frac{dy(\xi,1-\xi)}{d\xi}\right)^2} d\xi. \tag{D.37}
$$

Assembly of the elemental matrices and vectors yields the following finite dimensional problem formulation:

$$
(\mathbf{K} + \mathbf{M} + \mathbf{P}_\partial)\mathbf{c} = \mathbf{f}, \tag{D.38}
$$

where $\mathbf{c} = [c_1, c_2, \cdots, c_N]^T$. We note that the ellipticity of the sesquilinear forms $b(\psi, g)$ and $b^*(\psi, g^*)$ carries over to the finite-dimensional subspace (for conforming finite elements) [47]. As a result, the finite element matrix $(\mathbf{K}+\mathbf{M}+\mathbf{P}_\partial)$ is positive definite and thus has a bounded inverse.