ADAPTIVE FINITE ELEMENT METHODS FOR FLUORESCENCE DIFFUSE OPTICAL TOMOGRAPHY

By

Lu Zhou

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 ENGINEERING

Approved by the
Examining Committee:

Birsen Yazıcı, Thesis Adviser

Gary J. Saulnier, Member

Lester A. Gerhardt, Member

George Xu, Member

Rensselaer Polytechnic Institute
Troy, New York

June 2010
(For Graduation August 2010)

CONTENTS

APPENDICES

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGMENT

First and foremost, I thank my wife, Ruoxin Zhang, who gave up so much to be with me and started a new life on the other side of the earth, as well as my mother and father, who let their only son pursue his dream and can only see him every other year. I would like to dedicate this thesis to them and thank them for their unconditional love and support throughout my life.

I am very fortunate to have an exceptional advisor, Prof. Birsen Yazıcı. Four years ago, she discovered me from hundreds of applications and invited me to the Ph.D problem in the United States. Since then, she always trusted in me, encouraged and motivated me to pursue a Ph.D. degree. She always wanted me to do my very best and spent her precious time with me to achieve this goal. I am so grateful for her patience, guidance, help, and support during my Ph.D study.

I would like to thank my other committee members, Prof. Gary J. Saulnier, Prof. Lester A. Gerhardt, and Prof. George Xu for dedicating their precious time on my thesis work. I appreciate their indispensable advice to my thesis work.

I would like to thank my former colleague, Dr. Murat Güven, and my current colleague, An Jin, for their kindly helps in my research. It has been a great pleasure working with them during my past Ph.D. years.

Finally, I would like to thank all my friends and my families, no matter where are they or where I am, I am so proud to have them in my life.

ABSTRACT

Fluorescence Diffuse Optical Tomography (FDOT) is an emerging molecular imaging modality with applications in small animal and deep tissue imaging. FDOT uses visible or near infrared light to reconstruct the concentration and pharmacokinetics, as well as the life time of fluorophores injected into the tissue, based on a mathematical model of light propagation in turbid media. Due to the diffusive nature of light propagation in tissue, FDOT image reconstruction is a highly nonlinear, inherently three-dimensional, and computationally intense inverse problem. This thesis focuses on developing discretization error analysis and subsequent spatially varying resolution techniques to address the tradeoff between the reconstruction accuracy and the computational requirements of FDOT.

In the first part of the thesis, we formulate the FDOT inverse problem as an optimization problem with Tikhonov regularization under the assumption of noise-free measurements. We next analyze the effect of forward and inverse problem discretizations on the accuracy of FDOT reconstruction. Our analysis identifies several factors that determine the extent to which the discretization affects the accuracy of reconstructed fluorescence optical images. Based on our error analysis, we develop adaptive mesh generation algorithms with the objective of increasing the reconstruction accuracy while keeping the discretized forward and inverse problem sizes within allowable limits. In the simulation study, we demonstrate the effectiveness of our new algorithms and compare it with those of the uniform and conventional adaptive meshing schemes.

In the second part of the thesis, we consider measurements corrupted by additive noise and formulate the FDOT inverse problem as an optimization problem in the maximum *a posteriori* framework. We analyze the effect of measurement noise in the FDOT forward and inverse problem discretizations and develop adaptive mesh generation alogrithms that take into account noise statistics as well as *a priori* information on the fluorophore concentration. In the simulation study, we evaluate the performance of our new adaptive mesh generation algorithms and compare their performance with those of the uniform meshing scheme and the algorithms developed in the first part of the thesis.

We apply our new adaptive mesh generation algorithms to FDOT reconstruction using data from a phantom experiment, and demonstrate the practical advantages of our algorithms in real FDOT reconstruction.

Finally, we note that while our focus has been the FDOT inverse problem, the methods and algorithms developed in this thesis can be adapted to other partial differential equation based inverse parameter estimation problems, such as diffuse optical tomography, bioluminescence tomography, electrical impedance tomography, and microwave tomography.

CHAPTER 1
INTRODUCTION

1.1 DOT and FDOT Imaging

Diffuse Optical Tomography (DOT) is a promising medical imaging technique [1, 2], that uses visible or Near Infrared (NIR) light to image the optical properties inside biological tissue. Light, when traveling inside the tissue, is either elastically scattered or absorbed by three major types of molecules: water, oxygenated hemoglobin (HbO), and deoxygenated hemoglobin (Hb). These three types of molecules show a relatively weak absorption to the light with the wavelength ranging from 700 nm to 1000 nm [3]. The light with wavelength in this "spectral window" can penetrate the tissue up to several centimeters, and can be used to explore and recover the absorption as well as the scattering properties of the tissue in terms of the spatially varying absorption and scattering coefficient images. Furthermore, due to different absorption properties of these three types of molecules at different wavelengths, their respective concentrations can also be recovered using the spectroscopy technique [4]. In this respect, DOT offers an opportunity to image the hemoglobin concentration and oxygen saturation, which reflect the metabolic level and functional information of the tissue [4].

DOT has a number of advantages as compared to the traditional anatomical imaging modalities, such as ultrasound, X-ray, Computed Tomography (CT), and Magnetic Resonance Imaging (MRI). First, DOT is a non-invasive and non-ionizing imaging modality. Second, the instrumentation required for DOT is relatively less expensive and can be made portable. Third, DOT provides functional information that is not available from those anatomical imaging modalities. With these advantages, some of the potential applications of DOT include optical mammography [5–10], cognitive activity monitoring [11–13], and neonatal brain imaging [14].

Fluorescence Diffuse Optical tomography (FDOT) is an emerging molecular imaging technique for small animal or deep tissue imaging based on the same light propagation model as the one used in DOT. FDOT provides a method to visualize and quantify some specific molecular targets, and chemical and biological processes *in vivo* [15, 16]. In FDOT imaging, fluorescent probes, also referred to as "fluorophore", can be developed to target on specific genes and proteins. These fluorescent probes are either molecules [17] or small particles [18] that can absorb the energy from an external light source and re-emit light at a lower-energy level with longer wavelength. By illuminating the tissue and measuring the emission light coming out, one can use tomographic methods similar to that used in DOT to retrieve the fluorophore concentration and lifetime (the time in which the emitting light intensity decays to $1/e$ of the its initial intensity) information. The concentration information of fluorophore directly indicates the distribution of those specific proteins or genes in tissue and how it changes over time [19]. Additionally, fluorescence lifetime can also provide valuable information on tissue oxygenation, pH, and glucose concentration [20].

Some traditional radiologic imaging techniques, such as MRI, CT, Single Photon Emission Computed Tomography (SPECT), and Positron Emission Tomography (PET), are also adapted and used in the field of molecular imaging. FDOT has the similar advantages over these traditional techniques as DOT. Also similar to DOT, FDOT can be classified into three categories based on the employed method of illumination and detection [16]: Continuous Wave (CW) [21–24], Frequency Domain (FD) [25–27], and Time Domain (TD) [28–30] methods. CW method uses light with constant intensity generated and measured by simple low-cost optical equipments. This method has the advantage of comparatively higher Signal-to-Noise Ratio (SNR) than those of FD and TD methods, because of the stability and better noise characteristic of the optical components [16]. However, it can not recover the lifetime information of fluorophore. FD method uses light with modulated intensity, and the measurements of light intensity and phase shift at detector positions can be used to obtain both localization and lifetime information of fluorophore. On the other hand, TD method, that uses fast light pulses to illuminate tissue and measures the time-resolved light intensity at detector positions, can also be used to retrieve the lifetime information. One advantage of TD method is its higher resolution achieved by using the information contained in less diffusive early arriving photons [31]. However both FD and TD methods require more complex and expensive optical equipments as compared to that used by the CW method, and their SNRs are both lower than the CW method [16].

In recent years, researchers began to use direct-lens-coupled camera in FDOT data collection to replace the traditionally used light-guiding fibers. This approach offers high spatial sampling of the measurement and increases the number of total measurements from traditionally 10^2-10^3 to 10^4-10^6 [16]. This method was reported to be able to improve the reconstruction resolution to sub-millimeter level [22]. Further, a non-contact imaging technique [24, 32] using the similar data collection method with camera was introduced into FDOT field. This method eliminates the use of matching fluid, and instead a surface capture device is used to obtain the tissue boundary information [32]. With a model that describes the combined light propagation from tissue to air [33], this method provides the possibility of using FDOT to explore arbitrary bounded tissue or animal, greatly simplifies the experiment process, and improves the flexibility of optical imaging technique [24].

1.2 Forward and Inverse Problems of FDOT

In this thesis, we focus on the image reconstruction aspect of FDOT imaging. In this context, we consider FDOT imaging as an inverse problem that involves recovering the unknown optical properties (i.e., the fluorophore concentration and lifetime inside tissue) from the boundary measurements of light based on a forward model of light propagation. The FDOT forward problem involves computing the light field in tissue given the optical properties of the medium based on the FDOT forward model. In this section, we give a brief introduction to the forward and inverse problems of FDOT from a computational perspective.

There are two types of mathematical models available to describe light propagation in biological tissue: The Radiative Transport Equation (RTE) and the diffusion equation. The radiative transport equation [34], also referred to as Boltzmann transport equation, treats the light as a large number of photons propagating in the tissue, and it is based on the conservation law of the radiance. Although the radiative transport equation is considered as a more accurate model than the diffusion equation model, its solution is very difficult to obtain either analytically or numerically due to its integro-differential structure [35]. Under the assumption that the tissue has much stronger scattering property than the absorption property in the NIR range, the radiative transport equation can be approximated by the diffusion equation [36].

In FDOT, a pair of coupled diffusion equations are used to model the excitation and emission lights [37, 38]:

$$
\begin{aligned}
-\nabla \cdot D_x(\boldsymbol{r})\nabla\phi_x(\boldsymbol{r},t) + \mu_{ax}(\boldsymbol{r})\phi_x(\boldsymbol{r},t) + \frac{1}{c}\frac{\partial\phi_x(\boldsymbol{r},t)}{\partial t} &= S(\boldsymbol{r},t), \qquad (1.1)\\
-\nabla \cdot D_m(\boldsymbol{r})\nabla\phi_m(\boldsymbol{r},t) + \mu_{am}(\boldsymbol{r})\phi_m(\boldsymbol{r},t) + \frac{1}{c}\frac{\partial\phi_m(\boldsymbol{r},t)}{\partial t} &= \frac{1}{\tau}\eta\mu_{axf}(\boldsymbol{r})\\
&\quad \cdot \int_0^t \exp\left(\frac{t'-t}{\tau}\right)\phi_x(\boldsymbol{r},t')dt', \qquad (1.2)
\end{aligned}
$$

where subscripts x and m denote the excitation and emission, respectively, $\phi_{x,m}(\boldsymbol{r},t)$ is the light field at position $\boldsymbol{r}$ and at time t, $D_{x,m}(\boldsymbol{r})$ and $\mu_{ax,am}(\boldsymbol{r})$ are the diffusion coefficient and absorption coefficient of the tissue, μ_{axf}, τ, and η are the absorption coefficient, lifetime and quantum efficiency of the fluorophore, respectively. Since the absorption coefficient of the fluorophore is proportional to the fluorophore concentration, therefore retrieving the fluorophore concentration information is equivalent to recovering the absorption coefficient of the fluorophore. Although this approximation may fail in the tissue with stronger absorption than the scattering property [39, 40], it is valid in many applications that involve biological tissue [2].

There are three types of methods used to solve the forward problem defined based on the RTE or diffusion equation: Analytical method, statistical method, and numerical method. The analytical method attempts to find a closed form expression for the Green's function of the diffusion equations (1.1) and (1.2). However, analytical solution only exists for simple imaging domain geometries and homogeneous background optical properties [41–43]. The statistical method solves the forward problem based on the RTE by simulating a large number of photons and calculating the trajectory of each photon. One of the most widely used methods is the Monte Carlo method [44] which can solve the forward problem with very high accuracy. However, because of the intensive computational requirements of this method, it is typically used to solve the forward problem only where diffusion equation does not hold [44–46], and to validate the solutions obtained by other less accurate methods [47, 48]. The numerical method can overcome the analytical method's inability to solve for light field in arbitrary domain geometry and requires much fewer computational resources than the statistical method [49–51]. In this thesis, we model the FDOT forward problem by a pair of coupled diffusion equations in the frequency domain, and use one of the most widely used numerical methods: Finite Element Method (FEM) to solve the forward problem.

Let the forward problem of FDOT be represented by the coupled diffusion equations (1.1) and (1.2). Then, we can define a relationship between the light field, y, on the boundary of imaging domain and the spatially varying fluorophore concentration or lifetime, x, as

$$y = \mathcal{A}(x), \tag{1.3}$$

where $\mathcal{A}$ is a nonlinear operator determined by (1.1) and (1.2). Then, the inverse problem of FDOT involves solving x given y based on the relationship defined in (1.3). This is a highly nonlinear inverse problem. Furthermore, it is ill-posed in the following two senses:

- It does not have a unique solution, because the number of unknowns is usually much larger than the number of measurements.
- It does not have a stable solution, i.e., the solution is highly sensitive to small variations in the measurements.

To solve the FDOT inverse problem, both its nonlinearity and ill-posedness need to be appropriately addressed in the inverse problem formulation.

The nonlinearity of the FDOT inverse problem in (1.3) can be addressed by an iterative linearization. When the variation of the optical property is small with respect to a known background value (in FDOT, this means the absorption coefficient of the fluorophore is much smaller than that of the tissue or the lifetime is close to a known background value), the nonlinear relationship (1.3) can be linearized by a Born- or Rytov-type of approximation [52, 53]. After the discretization of the linearized relationship, (1.3) can be reduced to a matrix-vector equation:

$$\mathbf{y} = \boldsymbol{A}\mathbf{x}, \tag{1.4}$$

where $\mathbf{y}$ and $\mathbf{x}$ are the discretized measurements and fluorophore concentration or lifetime, and $\boldsymbol{A}$ is the matrix representation of the linearized operator $\mathcal{A}$, often referred to as the "weight matrix" [22].

There are many inversion methods that can be used to solve (1.4) while countering its ill-posedness. These methods include the truncated singular value decomposition, algebraic reconstruction technique, Krylov subspace method and conjugate gradient method [54, 55].

To address the stability and non-uniqueness of the FDOT inverse problem, (1.3) can be regularized [56, 57] as follows:

$$x = \arg_x \min \left\{ \|y - \mathcal{A}(x)\|^2 + \lambda \|\mathcal{L}(x, y)\|^2 \right\}, \tag{1.5}$$

where $\mathcal{L}(x, y)$ is a regularization functional that represents some *a priori* information about x and y, λ is the regularization parameter, and $\|\cdot\|$ is an appropriately chosen norm. The first term in (1.5) forces the solution x to fit to the measurements y, and the second term constrains x to prevent an unstable solution. In this respect, λ is used to balance these two terms and control the solution toward the measurements or *a priori* information. There are many nonlinear optimization methods that can be used to solve (1.5) including the Gauss-Newton [58, 59] and iterative coordinate descent methods [57].

1.3 Adaptive FEMs for FDOT and Other Inverse Problems

FDOT poses a computationally intense imaging problem. This stems from the necessity of numerically solving the interdependent forward and inverse problems. If the nonlinear inverse problem formulation is used in FDOT reconstruction, one needs to linearize and solve the inverse problem at many iteration steps until it converges. At each iteration, solving the inverse problem also requires first solving the forward problem, comprised of a pair of coupled diffusion equations, for each source or detector. For both forward and inverse problems, numerical solutions are merely approximations of the exact solutions, and they possess error as a result of the problem discretization. In this respect, the problem discrtization presents a tradeoff between the accuracy and the computational complexity of FDOT reconstruction. One way to achieve higher accuracy is to reduce the size of the mesh used to discretize the problem. However, reducing the size of the mesh implies a substantial increase of the discretization points due to the three-dimensional nature of the FDOT imaging problem, thereby increasing the resulting discretized problem size as well as the computational complexity of the image reconstruction.

One way to address this tradeoff, is to use the adaptive FEMs to solve both FDOT forward and inverse problems. The adaptive FEMs use adaptive meshes with spatially varying resolution for the problem discretization. With a constraint on the computational complexity, the adaptive meshes can produce more accurate numerical solutions than those produced by the meshes with uniform resolution. To generate the meshes that can effectively reduce the discretization error, the adaptive mesh generation algorithm must be designed with a solid understanding about the effects of the forward and inverse problem discretizations on the accuracy of FDOT reconstruction. Due to the interdependence of the FDOT forward and inverse problems, the discretizations of both problems affect the final reconstruction accuracy in a complex way. While the effect of inverse problem discretization can be deduced rather intuitively [60], the discretization error in the forward problem solutions can also result in unexpected artifacts in optical imaging [61]. In this case, the adaptive meshes refined independently for the forward and inverse problems may not be as effective in reducing the error due to discretization as those refined by taking the interdependence of two problems into account [60]. Furthermore, in the imaging problem, the measurement noise is also an important factor that may interact with the discretization to affect the accuracy of the image reconstruction. Therefore, adaptive mesh generation algorithms designed for FDOT imaging also need to take into account the impact of measurement noise.

There is a vast degree of work on the estimation and analysis of discretization error in the numerical solutions of Partial Differential Equations (PDEs) solved by

FEMs [62–67]. However, in the area of inverse problems, which often present as parameter estimation problems governed by PDEs, relatively little has been published. This parameter estimation problem is often formulated as an optimization problem: Given some PDEs, find the values of some parameters in these PDEs that can minimize the difference between the actual observation and the prediction calculated based on them [68]. Several error estimates for this type of problem have been derived (see [69–73]), based on which adaptive meshing methods can be developed to reduce the error due to discretization. For applications of these works (see, for example [74, 75]), Beilina *et al.* derived an *a posteriori* error estimate and developed an adaptive meshing method for the solution of an inverse acoustic scattering problem.

In the area of DOT, Arridge *et al.* [61] numerically showed that the error resulting from the forward problem discretization can lead to significant artifacts in the reconstructed optical images. In that work, these artifacts are minimized by modeling the discretization error in the forward problem solution as a random variable and reconstructing the image within the Bayesian framework. Furthermore, the adaptive meshing methods in FEMs field were adapted to DOT imaging to improve the accuracy and computational efficiency of the image reconstruction. In [76], Gu *et al.* developed a dual meshing method for DOT. This method uses two meshes with different discretization levels in DOT reconstruction: fine mesh for the forward problem and coarse mesh for the inverse problem. Furthermore, they proposed an adaptive meshing method similar to those in [77, 78], which adaptively refines the mesh around the region of heterogeneity, identified either from the co-registered ultrasound image [77] or the previous DOT reconstruction [76, 78]. In [79], Guven *et al.* presented a rigorous approach to analyze the effect of discretization on the accuracy of DOT reconstruction, which further led to the development of a new adaptive mesh generation method for DOT [60]. Their new method takes into account the interdependence of forward and inverse problems when discreizatizing either of them and can effectively reduce the error due to discretization in the reconstructed optical images.

In the area of FDOT, it has been shown that, when the computational resources are constrained, using the meshes with spatially varying resolution can result in higher accuracy and resolution in the reconstructed images than using the meshes with uniform resolution [59, 80, 81]. In [59, 80], Bangerth *et al.* formulated the FDOT image reconstruction problem as a PDE-constrained optimization problem, and employed a mesh refinement criteria suggested in a dual weighted residual framework [69]. In [82], to achieve fast and robust parameter mapping between the adaptively refined/derefined meshes of forward and inverse problems, Lee *et al.* developed an algorithm to identify and resolve the intersections of tetrahedral finite elements. In [81], this algorithm was utilized in FDOT reconstruction where the meshes for the forward and inverse problems are independently refined based on an *a posteriori* error estimate.

Adaptive meshing algorithms have been proposed for both DOT and FDOT imaging before [76, 77, 77, 78, 81]. However, in all these studies, existing adaptive meshing methods/algorithms are directly applied to the DOT and FDOT imaging without giving any specific consideration to problems in hand. As such these algorithms are not expected to produce optimal results in addressing the computational complexity vs. accuracy tradeoff as demonstrated in our simulation study. There are a number of studies, including our previous study [59, 60, 79], that take into account forward and inverse problem interdependency in adaptive meshing. However, these studies do not take into account the effect of measurement noise in designing adaptive meshing algorithms. In Chapter 5, we describe an adaptive mesh generation algorithm that takes into account the statistics of measurement as well as the interdependency of the FDOT forward and inverse problems. We demonstrate the effectiveness of our method in simulation and phantom studies over existing approaches.

1.4 Thesis Objectives

In this thesis, we focus on analyzing the effect of forward and inverse problem discretizations on the accuracy of FDOT reconstruction and developing adaptive mesh generation algorithms to address this tradeoff. More specifically, in our analysis, we determine the factors that affect the reconstruction accuracy due to forward and inverse problem discretizations. Based on our analysis, we develop the adaptive mesh generation algorithms for FDOT forward and inverse problems. These algorithms generate the meshes with spatially varying resolution, and can be used to reduce the error due to discretization in FDOT reconstruction.

The objectives of this thesis can be summarized as follows:

- To define a figure of merit that quantitatively evaluates the accuracy of the reconstructed optical image, and show explicitly the effect of the forward and inverse problem discretizations on the reconstruction accuracy. In this respect, we identify the factors that can affect the reconstruction accuracy together with the forward and inverse problem discretizations (i.e., the choice of basis functions and discretization points).
- To develop adaptive mesh generation algorithms for the FDOT forward and inverse problems that take into account the factors identified in the discretization effect analysis to improve accuracy and resolution of FDOT imaging while keeping the computational requirements within allowable limits. We develop several adaptive mesh generation algorithms that can address the tradeoff between the computational complexity and the reconstruction accuracy of FDOT imaging.
- To compare our adaptive mesh generation algorithms with uniform and conventional adaptive meshing methods in FDOT reconstruction, and demonstrate their practical advantages.

To achieve the objectives stated above, our work in this thesis was developed in two parts: In Chapters 2 and 3, we first analyze the effect of discretization on the reconstruction accuracy and develop adaptive mesh generation algorithms for FDOT under the assumption that the measurements are noise-free. This assumption simplifies our analysis and allows us to identify the key factors related to the imaging geometry and heterogeneity involved in the forward and inverse problem discretizations, that effect the accuracy of the FDOT reconstruction. In Chapters 4 and 5, we address the impact of measurement noise in FDOT forward and inverse problem discretizations and its effect in reconstruction accuracy.

The major contributions of this thesis can be summarized as follows:

- The discretization error analysis in the noise-free case presents a clear relationship among the accuracy of FDOT reconstruction, factors associated with the FDOT imaging problem (such as source-detector configuration, their positions with respective to fluorophore heterogeneity and the inverse problem formulation), and forward and inverse problem discretizations. To the best of our knowledge, this part of the thesis is the first study that provides a comprehensive analysis on the error due to discretization in FDOT imaging literature. This analysis further suggests novel adaptive mesh generation algorithms for FDOT imaging to address the tradeoff between the reconstruction accuracy and computational complexity of FDOT imaging.

- The discretization error analysis in the presence of measurement noise takes the noise statistics as well as the *a priori* information on the fluorophore concentration into account, and shows how the discretization together with the noise and *a priori* information affect the accuracy of FDOT reconstruction. To the best of our knowledge, this part of the thesis is the first study addressing the discretization error and adaptive mesh generation in the presence of measurement noise in the FDOT, as well as the DOT imaging literature.

- The adaptive mesh generation algorithms developed in this thesis can significantly improve the FDOT reconstruction accuracy while constraining the computational complexity within allowable limit, as compared to the uniform and conventional adaptive meshing schemes.

- The discretization error analysis framework and the adaptive mesh generation algorithms developed in this thesis can also be adapted to other inverse parameter estimation problems involving PDEs, such as bioluminescence tomography, electrical impedance tomography, and microwave tomography.

1.5 Thesis Outline

Throughout this thesis, we focus on the estimation of the fluorophore concentration in a bounded optical domain using the CW method. In this respect, for the forward problem of FDOT, we model the light propagation in the bounded domain by a pair of coupled frequency-domain diffusion equations with the Robin-type boundary conditions. After transforming this boundary value problem into a variational problem, we use the Galerkin projection method with first-order Lagrange basis functions to discretize and solve the forward problem. For the inverse problem of FDOT, we consider an iterative linearization method based on the Born approximation to address the nonlinearity. The outline of the thesis is as follows:

In Chapter 2, we analyze the effect of discretization on FDOT imaging when the measurements are noise-free. Thus, we assume that there is no error in the reconstructed images due to noise. In this case, we assess the reconstruction accuracy using the $H^1(\Omega)$ norm of the error due to discretization in the reconstructed image. We formulate the inverse problem as an optimization problem at each iteration of the linearization with zeroth- and first-order Tikhonov regularization. Next, with an appropriate boundary condition, we convert the optimization problem into a variational problem, and discretize and solve it using the Galerkin projection method with first-order Lagrange basis functions. Following the discretizations, we derive two upper bounds for the error in the reconstructed image due to the forward and inverse problem discretizations, respectively. Unlike the conventional error bounds in the finite element theory, these two new error bounds take into account the mutual dependence of the forward and inverse problems and identify the specific factors in the FDOT imaging problem. These factors include the number of sources and detectors, their configuration and their positions with respect to the fluorophore heterogeneity, and the formulation of the inverse problem.

In Chapter 3, based on the two error bounds derived in Chapter 2, we develop two new adaptive mesh generation algorithms for the forward and inverse problems of FDOT for the noise-free case. These two new adaptive mesh generation algorithms provide improved reconstruction accuracy as compared to the conventional adaptive and uniform meshing schemes while keeping the size of the discretized problems under the allowable limits. In contrast to the adaptive meshing method suggested by the conventional error bound in finite element theory, our new algorithms address the interdependence between the solutions of the forward and inverse problems, and take into account the source and detector configuration, and their relative locations with respect to the fluorophore heterogeneity, as well as the inverse problem formulation and the regularization parameters. After describing each algorithm in detail, we also illustrate their mesh refinement process through a toy problem, compare them with the conventional method and briefly discuss their computational complexities. At the end of this chapter, we present a series of three-dimensional simulation studies to illustrate the practical advantages of our new adaptive mesh generation algorithms in FDOT reconstruction. Our numerical simulations show the improvements, resulting from our new algorithms, in terms of accuracy, resolution, and small target detectability of the reconstructed optical images, as well as the convergence rate of the discretized inverse solution to the exact inverse problem solution.

In Chapter 4, we analyze the effect of discretization on FDOT imaging when the measurements are corrupted by additive noise. We assume that the measurement noise is an independent, zero-mean Gaussian process. Similarly, we assume that the fluorophore concentration is a Gaussian random field with known first- and second-order statistics. In this respect, we formulate the FDOT inverse problem in the Bayesian framework and consider the *Maximum A Posteriori* (MAP) estimator of the fluorophore concentration taking into account *a priori* information on fluorophore concentration at each linearization step. Following the approach presented in Chapter 2, we transform these optimization problems into variational problems and discretize them with the Galerkin projection method. Since the measurements are contaminated with random noise, the estimate of fluorophore concentration as well as the error due to discretization are random fields rather than the deterministic functions as considered in Chapter 2. To quantitatively assess the error due to discretization, we define the Mean-Square-Error (MSE) between the exact solution and the discretized solution of the inverse problem, and use this as the figure of merit to evaluate the reconstruction accuracy. Then, we analyze the effect of discretization and present two new upper bounds on the MSE that present a direct relationship among the MSE due to discretization, the forward and inverse problem discretizations, the measurement noise, and the *a priori* information on the fluorophore concentration. Unlike the upper bounds given in Chapter 2, the new error bounds not only include the factors (such as the number of sources and detectors, their configuration and their positions with respect to the fluorophore heterogeneity) identified by the upper bounds in Chapter 2, but also take into account the noise statistics as well as the *a priori* information on fluorophore concentration.

In Chapter 5, based on the two upper bounds derived in Chapter 4, we develop two new adaptive mesh generation algorithms for the forward and inverse problems of FDOT, taking into account the noise statistics and *a priori* information on fluorophore concentration. Additionally, unlike the algorithms in Chapter 3, these two new algorithms aim to improve the reconstruction accuracy in the mean square sense, while keeping the size of the discretized problems under an allowable limit. After describing each algorithm in detail, we address several practical issues in implementing the algorithms, briefly discuss their computational complexities, and compare them with the ones in Chapter 3 and the conventional adaptive meshing method. At the end of this chapter, we present a series of three-dimensional simulation studies in the presence of measurement noise to demonstrate the practical advantages of our new algorithms over the ones in Chapter 3. The simulation results show that, in the presence of measurement noise, our new algorithms can more effectively reduce the MSE and improve the convergence rate of the discretized inverse problem solution to the exact inverse problem solution, than the algorithms given in Chapter 3.

In Chapter 6, we used data from a phantom experiment to demonstrate the performance of our adaptive mesh generation algorithms in the real FDOT reconstruction. Since the measurements are always corrupted by noise in the real FDOT reconstruction, we apply our adaptive mesh generation algorithms, proposed in Chapter 5, to FDOT reconstruction using data obtained from a phantom experiment, and compare our algorithm with the uniform and conventional adaptive meshing schemes. The phantom experiment results demonstrate the effectiveness of our algorithms in real data.

In Chapter 7, we conclude the thesis and provide a discussion on the potential future work.

We summarize the performance of conventional adaptive meshing algorithm and the proposed adaptive mesh generation algorithms in this thesis in Table 1.1.

Table 1.1: Comparison of different adaptive mesh generation algorithms.

	Conventional adaptive meshing algorithm	Algorithms in Chapter 3	Algorithms in Chapter 5
H^1 norm error reduction	0-15%	50-60%	N/A
MSE reduction	N/A	25-40%	70-75%
Resolution	8.5 mm	6.5 mm	6.5 mm
Computational complexity[1]	$\mathcal{O}(N_\Delta)$	$\mathcal{O}(N_\Delta N_D)$ $\mathcal{O}(N_\Delta N_S)$ $\mathcal{O}(N_\Delta N_S N_D)$	$\mathcal{O}(N_\Delta N_D)$ $\mathcal{O}(N_\Delta N_S)$ $\mathcal{O}(N_\Delta N_S N_D)$

---

[1] N_S is the number of sources, N_D is the number of detectors, N_Δ is the number of element in the mesh used for problem discretization. For algorithms in Chapters 3 and 5, from top to bottom, the computation complexities are given for refining the mesh for a single source, a single detector and inverse problem.

CHAPTER 2
EFFECT OF DISCRETIZATION IN FDOT IN THE NOISE-FREE SITUATION[1]

2.1 Introduction

In this chapter, we analyze the effect of discretization on the accuracy of FDOT imaging when the measurements are noise-free. In this respect, there is no error due to the noise in the reconstructed image. This assumption simplifies our analysis and allows us to identify the key factors specific to the imaging problem that show how the forward and inverse problem discretizations impact the accuracy of reconstruction. In particular, we focus on the estimation of the fluorophore concentration in a bounded optical domain using the CW method. For the forward problem, the light propagation in this domain is modeled by a pair of coupled frequency-domain diffusion equations with the Robin-type boundary conditions. After transforming this boundary value problem into a variational problem, we use the Galerkin projection method with first-order Lagrange basis functions to discretize and solve the forward problem. For the inverse problem, we consider an iterative linearization method based on the Born approximation to address the nonlinearity. At each linearized iteration, we formulate the inverse problem as an optimization problem that enables us to incorporate *a priori* information in the form of zeroth- and first-order Tikhonov regularization. Then, with the appropriate boundary condition, we convert this optimization problem into a variational problem, which can be discretized and solved by the Galerkin method as we use in the forward problem.

Following the discretizations, we derive two upper bounds for the error in the reconstructed image due to discretization the forward and inverse problem discretizations, respectively. Unlike the conventional error bounds in the finite element theory, these two new error bounds take into account the mutual dependence of

---

[1] Portions of this chapter previously appeared as: M. Guven, L. Reilly-Raska, L. Zhou, and B. Yazıcı. Discretization error analysis and adaptive meshing algorithms for fluorescence diffuse optical tomography: part I. *IEEE Transactions on Medical Imaging*, 29(2):217-229, 2010.

the forward and inverse problems and identify the specific factors that may lead to high error in the reconstructed optical image. These factors include the number of sources and detectors, their configuration and their positions with respect to the fluorophore heterogeneity, and the formulation of the inverse problem. These new error bounds motivate the development of the new adaptive mesh generation algorithms, which are further discussed in the next chapter.

2.2 Notational Conventions

Throughout the thesis, we use capital cursive letters ($\mathcal{A}$) for operators and bold capital letters ($\boldsymbol{A}$) for matrices. We denote functions by lowercase letters (g and ϕ etc.) and their finite-dimensional approximations by corresponding uppercase letters (G and Φ etc.). We use bold to denote vectorized quantities such as $\boldsymbol{r}$ and $\mathbf{\Gamma}$. Table. 2.1 provides a summary of key variables, function spaces, and norms used throughout the thesis.

Table 2.1: Definition of function spaces and norms.

Notation	Explanation
$C(\Omega)$	Space of continuous functions on Ω
$L^\infty(\Omega)$	$L^\infty(\Omega) = \{f \mid \operatorname{ess\,sup}_\Omega \lvert f(\mathbf{x})\rvert < \infty\ \}$
$L^p(\Omega)$	$L^p(\Omega) = \{f \mid\ (\int_\Omega \lvert f(\mathbf{x})\rvert^p d\mathbf{x})^{1/p} < \infty\ \}$, $p \in [1, \infty)$
$H^p(\Omega)$	$H^p(\Omega) = \{f \mid\ (\sum_{\lvert z\rvert \leq p} \lVert D_w^z f\rVert_0^2)^{1/2} < \infty\ \}$, $p \in [1, \infty)$
$\lVert f\rVert_0$	The $L^2(\Omega)$ norm of f
$\lVert f\rVert_p$	The $H^p(\Omega)$ norm of f
$\lVert f\rVert_{p*}$	The dual norm of f in the dual space $H^{p*}(\Omega)$
$\lVert f\rVert_\infty$	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

2.3 Forward Problem of FDOT

We start with using a pair of coupled frequency-domain diffusion equations to model the light propagation in a fluorescent medium of a bounded domain $\Omega \subset R^3$

with Lipschitz boundary $\partial\Omega$ [37,38]:

$$
\begin{aligned}
-\nabla \cdot D_x(\boldsymbol{r})\nabla\phi_x(\boldsymbol{r},\omega) + \left(\mu_{ax}(\boldsymbol{r}) + \frac{j\omega}{c}\right)\phi_x(\boldsymbol{r},\omega) &= S_i(\boldsymbol{r},\omega), && (2.1)\\
-\nabla \cdot D_m(\boldsymbol{r})\nabla\phi_m(\boldsymbol{r},\omega) + \left(\mu_{am}(\boldsymbol{r}) + \frac{j\omega}{c}\right)\phi_m(\boldsymbol{r},\omega) &= \phi_x(\boldsymbol{r},\omega)\eta\mu_{axf}(\boldsymbol{r})\\
&\quad \cdot\frac{1-j\omega\tau(\boldsymbol{r})}{1+[\omega\tau(\boldsymbol{r})]^2}, && (2.2)
\end{aligned}
$$

where $\boldsymbol{r} = [r_1, r_2, r_3] \in \Omega$, ω is the source operating frequency, subscripts x, m denote the excitation and emission wavelengths, $\phi_{x,m}$ represents the light fields, $\mu_{ax,am}$ represents the absorption coefficient of the medium, $D_{x,m}$ represents the isotropic diffusion coefficients, μ_{axf} and τ are the absorption coefficient and the lifetime of the fluorophore, η is the quantum efficiency, and S_i is the i^{th} excitation source, $i = 1, \cdots, N_S$, where N_S is the number of sources. We assume that the diffusion coefficients are known and they are identical at both excitation and emission wavelengths in the closed domain, which implies $D(\boldsymbol{r}) := D_x(\boldsymbol{r}) = D_m(\boldsymbol{r})$, $\boldsymbol{r} \in \Omega \bigcup \partial\Omega$. Typically, we have following relationship among μ_{ax}, μ_{am}, and μ_{axf} for $\boldsymbol{r} \in \Omega$:

$$\mu_{ax}(\boldsymbol{r}) = \mu_{axe}(\boldsymbol{r}) + \mu_{axf}(\boldsymbol{r}), \tag{2.3}$$

$$\mu_{am}(\boldsymbol{r}) = \mu_{ame}(\boldsymbol{r}) + \mu_{amf}(\boldsymbol{r}), \tag{2.4}$$

where the subscript e denotes endogenous properties and f denotes exogenous properties. Without loss of generality, we assume that both $D(\boldsymbol{r})$ and $\mu_{ax,m}(\boldsymbol{r})$ are non-negative and bounded on Ω.

Since we consider using the CW method to estimate the fluorophore concentration in the medium, the frequency $\omega = 0$ in (2.1) and (2.2). We note that subsequent developments can be extended to include multiple frequencies in FD method of FDOT when τ is known. In this respect, we model the sources by a Gaussian function centered at source position $\boldsymbol{r}_i$ with angular frequency $\omega = 0$. Therefore, the NIR light propagation at the excitation and emission wavelengths can be model by the following boundary value problems:

$$\begin{aligned} -\nabla \cdot D(\boldsymbol{r})\nabla\phi_x(\boldsymbol{r},\boldsymbol{r}_i) + \mu_{ax}(\boldsymbol{r})\phi_x(\boldsymbol{r},\boldsymbol{r}_i) &= S_i(\boldsymbol{r}), \quad \boldsymbol{r} \in \Omega, &(2.5)\\ -\nabla \cdot D(\boldsymbol{r})\nabla\phi_m(\boldsymbol{r},\boldsymbol{r}_i) + \mu_{am}(\boldsymbol{r})\phi_m(\boldsymbol{r},\boldsymbol{r}_i) &= \phi_x(\boldsymbol{r},\boldsymbol{r}_i)\eta\mu_{axf}(\boldsymbol{r}), \quad \boldsymbol{r} \in \Omega, &(2.6)\end{aligned}$$

with the Robin-type boundary conditions [36]:

$$\begin{aligned} 2D(\boldsymbol{r})\frac{\partial\phi_x(\boldsymbol{r},\boldsymbol{r}_i)}{\partial n} + \rho\phi_x(\boldsymbol{r},\boldsymbol{r}_i) &= 0, \quad \boldsymbol{r} \in \partial\Omega, &(2.7)\\ 2D(\boldsymbol{r})\frac{\partial\phi_m(\boldsymbol{r},\boldsymbol{r}_i)}{\partial n} + \rho\phi_m(\boldsymbol{r},\boldsymbol{r}_i) &= 0, \quad \boldsymbol{r} \in \partial\Omega, &(2.8)\end{aligned}$$

where ρ is a parameter governing the internal reflection at the boundary $\partial\Omega$, and $\partial/\partial n$ denotes the directional derivative along the unit normal vector on the boundary.

In order to simplify the analysis later on, we make use of the adjoint problem associated with (2.6) and (2.8):

$$\begin{aligned} -\nabla \cdot D(\boldsymbol{r})\nabla g_m^*(\boldsymbol{r},\boldsymbol{r}_j) + \mu_{am}(\boldsymbol{r})g_m^*(\boldsymbol{r},\boldsymbol{r}_j) &= 0, \quad \boldsymbol{r} \in \Omega, &(2.9)\\ 2D(\boldsymbol{r})\frac{\partial g_m^*(\boldsymbol{r},\boldsymbol{r}_j)}{\partial n} + \rho g_m^*(\boldsymbol{r},\boldsymbol{r}_j) &= S_j^*(\boldsymbol{r}), \quad \boldsymbol{r} \in \partial\Omega, &(2.10)\end{aligned}$$

where $g_m^*(\boldsymbol{r},\boldsymbol{r}_j)$ is the solution of the adjoint problem for the j^{th} adjoint source S_j^* located at the detector position $\boldsymbol{r}_j \in \partial\Omega$, $j = 1, \cdots, N_D$, where N_D is the number of detectors. For a point adjoint source located at the detector position $\boldsymbol{r}_j$, $g_m^*(\boldsymbol{r},\boldsymbol{r}_j) = g_m(\boldsymbol{r}_j,\boldsymbol{r})$ holds where $g_m(\cdot,\boldsymbol{r})$ is the Green's function of (2.6) and (2.8) [36]. Note that in this thesis, we model the point adjoint source by a Gaussian function with sufficiently low variance. Then, the emission light field at $\boldsymbol{r}$ due to the source at $\boldsymbol{r}_i$ is given by the following nonlinear integral equation:

$$\phi_m(\boldsymbol{r},\boldsymbol{r}_i) = \int_\Omega g_m^*(\acute{\boldsymbol{r}},\boldsymbol{r})\phi_x(\acute{\boldsymbol{r}},\boldsymbol{r}_i)\eta\mu_{axf}(\acute{\boldsymbol{r}})d\acute{\boldsymbol{r}}. \tag{2.11}$$

Given N_S sources and N_D detectors, we define $\Gamma_{i,j}$ to be the measurement obtained by the j^{th} detector located at $\boldsymbol{r}_j$ due to the i^{th} source, then we have

$$\begin{aligned} \Gamma_{i,j} &:= \phi_m(\boldsymbol{r}_j, \boldsymbol{r}_i) \\ &= \int_\Omega g_m^*(\boldsymbol{r}, \boldsymbol{r}_j)\phi_x(\boldsymbol{r}, \boldsymbol{r}_i)\eta\mu_{axf}(\boldsymbol{r})d\boldsymbol{r}. \end{aligned} \tag{2.12}$$

The relationship between $\Gamma_{i,j}$ and μ_{axf} defined in (2.12) is nonlinear, because g_m^* is dependent on μ_{amf} nonlinearly which is in turn related to μ_{axf}, and the nonlinear dependence of ϕ_x on μ_{axf} is also clear. Further we note that, we assume that the measurements are noise-free in this chapter. This assumption allows us to eliminate the effect of noise in our error analysis, and to focus primarily on the effect of the discretization at the moment. In the next section, we will formally state the inverse problem, and address its nonlinearity and ill-posedness by the Born approximation and the Tikhonov regularization in an optimization framework respectively.

2.4 Inverse Problem of FDOT

Based on the forward problem formulation given in the previous section, the inverse problem can be defined as to recover the quantity μ_{axf} using the measurements $\Gamma_{i,j}$, $i = 1, \ldots, N_S$ and $j = 1, \ldots, N_D$, based on the nonlinear integral equation (2.12). To address the nonlinearity of this integral equation, we use the Born approximation [52, 53] and linearize (2.12) around a known background fluorophore absorption coefficient μ_{axf}^0: Let $\phi_x^0(\boldsymbol{r}, \boldsymbol{r}_i)$ and $g_m^{*,0}(\boldsymbol{r}, \boldsymbol{r}_j)$ be the solutions of (2.5)-(2.7) and (2.9)-(2.10) for $\mu_{axf} = \mu_{axf}^0$, then (2.12) can be linearized as

$$\Gamma_{i,j} = \int_\Omega g_m^{*,0}(\boldsymbol{r}, \boldsymbol{r}_j)\phi_x^0(\boldsymbol{r}, \boldsymbol{r}_i)\eta\mu_{axf}(\boldsymbol{r})d\boldsymbol{r}. \tag{2.13}$$

We let $g_j^*(\boldsymbol{r}) := g_m^{*,0}(\boldsymbol{r}, \boldsymbol{r}_j)$ and $\phi_i(\boldsymbol{r}) := \phi_x^0(\boldsymbol{r}, \boldsymbol{r}_i)$ suppressing the dependence of these functions on 0, x and m to simplify our notation. We also introduce $\mu(\boldsymbol{r}) := \eta\mu_{axf}(\boldsymbol{r})$ which we refer to as fluorophore concentration quantity to be reconstructed. Then (2.13) can be expressed as

$$\Gamma_{i,j} \;=\; \int_{\Omega} g_j^*(\boldsymbol{r})\phi_i(\boldsymbol{r})\mu(\boldsymbol{r})d\boldsymbol{r}. \tag{2.14}$$

Based on our linearized measurement model (2.14), we group the individual measurements into the following vector form:

$$\boldsymbol{\Gamma} := \left[\Gamma_{1,1}, \ldots, \Gamma_{1,N_D}, \Gamma_{2,1}, \ldots, \Gamma_{N_S,N_D}\right]^T, \tag{2.15}$$

and further define a vector-valued operator $\mathcal{A} : L^2(\Omega) \to \mathbb{R}^{N_S N_D}$ as

$$(\mathcal{A}\mu)_{ij} \;:=\; \int_{\Omega} a_{ij}(\boldsymbol{r})\mu(\boldsymbol{r})d\boldsymbol{r}, \tag{2.16}$$

where $a_{ij}(\boldsymbol{r}) := g_j^*(\boldsymbol{r})\phi_i(\boldsymbol{r})$. Combining (2.15) and (2.16), we write our measurement model as

$$\boldsymbol{\Gamma} = \mathcal{A}\mu. \tag{2.17}$$

We note that the Born approximation is valid and widely used in FDOT imaging when the perturbation of absorption coefficient is relatively small as compared to the known background absorption coefficient [38, 83]. However, when the perturbation of absorption coefficient is large, an iterative linearization method based on the Born approximation can be used to address the nonlinearity of the inverse problem. In this case the fluorophore concentration μ solved based on (2.17) can be iteratively refined based on the following model [84, 85]:

$$\boldsymbol{\Gamma} \;=\; \mathcal{A}^k\hat{\mu}_{k+1},$$

where $\hat{\mu}_{k+1}$ is the estimate of the fluorophore concentration at $(k+1)^{\text{th}}$ linearized iteration and

$$(\mathcal{A}^k\mu)_{ij} := \int_{\Omega} a_{ij}^k(\boldsymbol{r})\mu(\boldsymbol{r})d\boldsymbol{r},$$

where $a_{ij}^k = \phi_i^k(\boldsymbol{r})g_j^{*,k}(\boldsymbol{r})$, $\phi_i^k(\boldsymbol{r})$ and $g_j^{*,k}(\boldsymbol{r})$ are computed based on the fluorophore concentration $\hat{\mu}_k$ estimated at the k^{th} iteration.

Using the l^1 norm of the finite-dimensional space, an upper bound for the linear operator $\mathcal{A}$ can be given by

$$\|\mathcal{A}\|_{L^2(\Omega)\to l^1} \leq \sum_{i,j}^{N_S,N_D} \|g_j^* \phi_i\|_0. \tag{2.18}$$

Because of the boundedness and the finite-dimensional range of $\mathcal{A}$, we conclude that $\mathcal{A}$ is compact. Therefore the inverse problem based on (2.17) is ill-posed [86].

In this chapter, we address the ill-posedness of the FDOT inverse problem using the Tikhonov regularization in the optimization framework, which provides a suitable means for incorporating *a priori* information of the fluorophore concentration. In this respect, our inverse problem can be formulated by the following optimization problem where we seek a solution $\hat{\mu} \in H^1(\Omega)$ that minimize the objective functional J:

$$\begin{aligned} \hat{\mu} &= \min_{\mu \in H^1(\Omega)} J\left(\mu, \nabla\mu\right) \\ &= \min_{\mu \in H^1(\Omega)} \left[J_L(\mu) + J_R(\mu, \nabla\mu)\right], \end{aligned} \tag{2.19}$$

where the $H^1(\Omega)$ smoothness on the solution is imposed by the use of first-order Tikhonov regularization described in the following. In (2.19), the functional J_L that measures the difference between the predicted and actual measurements is given by

$$\begin{aligned} J_L(\mu) &= \|\mathbf{\Gamma} - \mathcal{A}\mu\|_{l^2}^2 \\ &= \sum_{i,j}^{N_S,N_D} \left[\Gamma_{i,j} - (\mathcal{A}\mu)_{i,j}\right]^2. \end{aligned} \tag{2.20}$$

The regularization term J_R regularize the ill-posed inverse problem by the *a priori* information of the fluorophore concentration. In this work, we assume that *a priori* information of μ and its gradient $\nabla\mu$ are available. Let $\beta_1(\boldsymbol{r})$ denote the *a priori* information on μ and $\boldsymbol{\beta}_2(\boldsymbol{r}) = [\beta_{21}(\boldsymbol{r}), \beta_{22}(\boldsymbol{r}), \beta_{23}(\boldsymbol{r})]^T$ denote the *a priori* information on $\nabla\mu$. We incorporate $\beta_1(\boldsymbol{r})$ and $\boldsymbol{\beta}_2(\boldsymbol{r})$ through zeroth- and first-order Tikhonov regularization terms as [87]:

$$J_R(\mu, \nabla\mu) = \lambda_1 \int_\Omega [\mu(\boldsymbol{r}) - \beta_1(\boldsymbol{r})]^2 \, d\boldsymbol{r} + \lambda_2 \int_\Omega |\nabla\mu(\boldsymbol{r}) - \boldsymbol{\beta}_2(\boldsymbol{r})|^2 d\boldsymbol{r}, \tag{2.21}$$

where λ_1, $\lambda_2 > 0$ are regularization parameters. There are a number of methods in choosing appropriate regularization parameters (see [88–92]). In this thesis, we assume that λ_1 and λ_2 are properly chosen.

Substituting (2.20) and (2.21) into (2.19), our inverse problem can be written explicitly as:

$$\begin{aligned} \hat{\mu} = \min_{\mu \in H^1(\Omega)} & \left[\sum_{i,j}^{N_S, N_D} [\Gamma_{i,j} - (\mathcal{A}\mu)_{i,j}]^2 + \lambda_1 \int_\Omega [\mu(\boldsymbol{r}) - \beta_1(\boldsymbol{r})]^2 \, d\boldsymbol{r} \right. \\ & \left. + \lambda_2 \int_\Omega |\nabla\mu(\boldsymbol{r}) - \boldsymbol{\beta}_2(\boldsymbol{r})|^2 \, d\boldsymbol{r} \right]. \end{aligned} \tag{2.22}$$

By taking the Gâteaux derivative of the objective functional in (2.22) and defining appropriate boundary conditions for it, we can convert this optimization problem into a boundary value problem. It can be shown that the solution of (2.22) satisfies $\partial J/\partial\mu(\mu, \nabla_q\mu) = 0$, where ∇_q is the gradient with respect to the r_q direction for $q = 1, 2, 3$. In particular, if $J = \int u(\boldsymbol{r}, \mu, \partial\mu/\partial r_q) d\boldsymbol{r}$, the Gâteaux derivative of J is defined by [93]

$$\frac{\partial J}{\partial \mu} = \frac{\partial u}{\partial \mu} - \sum_q \frac{\partial}{\partial r_q} \left(\frac{\partial u}{\partial \mu_q} \right). \tag{2.23}$$

Therefore, taking the Gâteaux derivative of (2.22) with respect to μ and setting it equal to zero yields:

$$(\mathcal{A}^*\mathcal{A}\mu)(\boldsymbol{r}) + \lambda_1\mu(\boldsymbol{r}) - \lambda_2\nabla^2\mu(\boldsymbol{r}) = f(\boldsymbol{r}), \tag{2.24}$$

where $f(\boldsymbol{r})$ containing the known terms from *a priori* information and measurement is defined as

$$f(\boldsymbol{r}) := (\mathcal{A}^*\boldsymbol{\Gamma})(\boldsymbol{r}) + \lambda_1\beta_1(\boldsymbol{r}) + \lambda_2 \sum_{q=1}^{3} \beta_{2q}(\boldsymbol{r}). \tag{2.25}$$

In both (2.24) and (2.25), $\mathcal{A}^* : \mathbb{R}^{N_S N_D} \to L^2(\Omega)$ is the adjoint operator of $\mathcal{A}$ defined as

$$(\mathcal{A}^*\mathbf{w})(\boldsymbol{r}) = [a_{11}(\boldsymbol{r}), \ldots, a_{1N_S}(\boldsymbol{r}), a_{21}(\boldsymbol{r}), \ldots, a_{N_S N_D}(\boldsymbol{r})] \cdot \mathbf{w}, \tag{2.26}$$

where $\mathbf{w} \in \mathbb{R}^{N_S N_D}$ and $a^*_{ij}(\boldsymbol{r}) = g^*_j(\boldsymbol{r})\phi_i(\boldsymbol{r})$, $i = 1, \ldots, N_S$ and $j = 1, \ldots, N_D$. We further define an operator $\mathcal{B} : L^2(\Omega) \to L^2(\Omega)$ as

$$\begin{aligned} (\mathcal{B}\mu)(\boldsymbol{r}) &:= (\mathcal{A}^*\mathcal{A}\mu)(\boldsymbol{r}) \\ &= \int_\Omega \sum_{i,j}^{N_S,N_D} a^*_{ij}(\boldsymbol{r}) a_{ij}(\acute{\boldsymbol{r}})\mu(\acute{\boldsymbol{r}})d\acute{\boldsymbol{r}}. \end{aligned} \tag{2.27}$$

Then (2.24) becomes

$$(\mathcal{B}\mu)(\boldsymbol{r}) + \lambda_1\mu(\boldsymbol{r}) - \lambda_2\nabla^2\mu(\boldsymbol{r}) = f(\boldsymbol{r}). \tag{2.28}$$

We consider (2.28) with the following Neumann boundary condition:

$$\frac{\partial\mu}{\partial n}(\boldsymbol{r}) = 0, \quad \boldsymbol{r} \in \partial\Omega, \tag{2.29}$$

where $\partial\mu/\partial n$ is the directional derivative of μ along the unit normal vector at the boundary $\partial\Omega$. The boundary condition in (2.29) implies that no changes in the fluorophore concentration occur across the boundary.

To apply the FEM on the boundary value problem (2.28) and (2.29), we need to first obtain the corresponding variational (weak) problem. Hence, we multiply both sides of (2.28) by a test function $\psi \in H^1(\Omega)$, and integrate it over Ω. Applying Green's first theorem to the last term on the left hand side and using the boundary condition in (2.29), we obtain

$$\int_\Omega \psi(\boldsymbol{r})\left[(\mathcal{B}\mu)(\boldsymbol{r}) + \lambda_1\mu(\boldsymbol{r})\right]d\boldsymbol{r} + \lambda_2\int_\Omega \nabla\psi(\boldsymbol{r}) \cdot \nabla\mu(\boldsymbol{r})d\boldsymbol{r} = \int_\Omega \psi(\boldsymbol{r})f(\boldsymbol{r})d\boldsymbol{r}. \tag{2.30}$$

A more convenient way to express (2.30) is using a bilinear form. Thus, we define

$$\begin{aligned} \mathcal{F}(\psi,\mu) &:= (\psi,\mathcal{B}\mu)+\lambda_1(\psi,\mu)+\lambda_2(\nabla\psi,\nabla\mu), && (2.31)\\ \mathcal{G}(\psi) &:= (\psi,f), && (2.32)\end{aligned}$$

where the inner product $(\cdot,\cdot)$ is defined by

$$\begin{aligned} (k,l) &:= \int_\Omega k(\boldsymbol{r})l(\boldsymbol{r})d\boldsymbol{r},\\ (\boldsymbol{k},\boldsymbol{l}) &:= \int_\Omega \boldsymbol{k}(\boldsymbol{r})\cdot\boldsymbol{l}(\boldsymbol{r})d\boldsymbol{r},\end{aligned}$$

for both scalar- and vector-valued functions. Then the variational form of the inverse problem can be expressed as

$$\mathcal{F}(\psi,\mu)=\mathcal{G}(\psi),\quad \forall\psi\in H^1(\Omega). \tag{2.33}$$

It can be shown that the bilinear form (2.31) is bounded and coercive for regularization parameters λ_1, $\lambda_2>0$ (see Appendix C). Therefore, by the Lax-Milgram lemma, a unique solution exists for the regularized inverse problem (2.33) for each pair of λ_1, $\lambda_2>0$ [86, 94]. In the next section, we will describe the discretization scheme used in this thesis for each of the separate forward and inverse problems as well as the combined forward and inverse problems.

2.5 Discretization of Forward and Inverse Problems

2.5.1 Forward Problem Discretization

We first express the coupled PDEs used to model the forward problem in their variational forms in order to apply FEM. We multiply both sides of (2.5) by a test function $\xi_1\in H^1(\Omega)$, and apply Green's theorem to the second derivative term. Then using the boundary condition in (2.7) we have

$$\int_\Omega(\nabla\xi_1\cdot D\nabla\phi_i+\mu_{ax}\xi_1\phi_i)d\boldsymbol{r}+\frac{1}{2\rho}\int_{\partial\Omega}\xi_1\phi_i dl \;=\; \int_\Omega \xi_1 S_i d\boldsymbol{r},\ \forall\xi_1\in H^1(\Omega). \tag{2.34}$$

It can be shown that a unique solution for (2.34) exists and is bounded [94], [95]. Similarly, for a test function $\xi_2 \in H^1(\Omega)$, the variational form for the adjoint forward problem (2.9) and (2.10) is

$$\int_\Omega \left(\nabla\xi_2 \cdot D\nabla g_j^* + \mu_{am}\xi_2 g_j^*\right) d\boldsymbol{r} + \frac{1}{2\rho}\int_{\partial\Omega} \xi_2 g_j^* dl = \frac{1}{2\rho}\int_{\partial\Omega} \xi_2 S_j^* dl,\ \forall \xi_2 \in H^1(\Omega), \tag{2.35}$$

for which it is possible to show that a unique bounded solution exists as well.

Let L_k denote the piecewise linear Lagrange basis functions. We define $Y_i^{N_i}(\Omega) \subset H^1(\Omega)$, $i = 1, \ldots, N_S$, as the finite-dimensional subspace spanned by $\{L_k, k = 1, \ldots, N_i\}$. Note that $\{L_k, k = 1, \ldots, N_i\}$ are associated with the set of points $\{\boldsymbol{r}_p, p = 1, \ldots, N_i\}$ on Ω. Similarly, we define $Y_j^{*,N_j}(\Omega) \subset H^1(\Omega)$, $j = 1, \ldots, N_D$, as the finite-dimensional subspace spanned by $\{L_k, k = 1, \ldots, N_j\}$, which are associated with the set of N_j points $\{\boldsymbol{r}_p, p = 1, \ldots, N_j\}$ on Ω.

Next, the functions ξ_1, ϕ_i in (2.5) and ξ_2, g_j^* in (2.35) are replaced by their finite-dimensional counterparts

$$\Xi_1^{N_i}(\boldsymbol{r}) := \sum_{k=1}^{N_i} p_k L_k(\boldsymbol{r}), \qquad \Phi_i^{N_i} := \textstyle\sum_{k=1}^{N_i} c_k L_k(\boldsymbol{r}), \tag{2.36}$$

$$\Xi_2^{N_j}(\boldsymbol{r}) := \sum_{k=1}^{N_j} p_k L_k(\boldsymbol{r}), \qquad G_j^{*,N_j} := \textstyle\sum_{k=1}^{N_j} d_k L_k(\boldsymbol{r}). \tag{2.37}$$

The representation $\Phi_i^{N_i}$ or G_j^{*,N_j} is an approximation to the function ϕ_i or g_j^* for each source or detector. This means that for each source and detector, the dimension of the solution can be different, and the parameters N_i and N_j may vary for each i and j, respectively. The finite-dimensional expansions are therefore dependent on the parameters N_i and N_j as represented by the superscript. However, we suppress this cumbersome notation as the dependence is clearly understood.

Substituting (2.36) and (2.37) into the variational forward problem (2.34) and

(2.35) yields the matrix equations

$$\mathbf{M}\mathbf{c}_i = \mathbf{q}_i, \tag{2.38}$$

$$\mathbf{M}^*\mathbf{d}_j^* = \mathbf{q}_j^*, \tag{2.39}$$

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

Further, let $\{\Omega_{ni}, n = 1, \ldots, N_\Delta^i\}$ denote the set of elements used to discretize (2.34), where N_Δ^i is the number of elements for the i^{th} source, such that $\bigcup_n^{N_\Delta^i} \Omega_{ni} = \Omega$ for all $i = 1, \ldots, N_S$. Similarly, let $\{\Omega_{mj}, m = 1, \ldots, N_\Delta^{*j}\}$ denote the set of elements used to discretize (2.35), where N_Δ^{*j} is the number of elements for the j^{th} detector, such that $\bigcup_m^{N_\Delta^{*j}} \Omega_{mj} = \Omega$ for all $j = 1, \ldots, N_D$.

The upper bounds for errors e_i and e_j^* defined as $e_i := \phi_i - \Phi_i$ and $e_j^* := g_j^* - G_j^*$ respectively on each finite element can be given by [94]:

$$\|e_i\|_{0,ni} \leq C\|\phi_i\|_{1,ni} h_{ni}, \tag{2.40}$$

$$\|e_j^*\|_{0,mj} \leq C\|g_j^*\|_{1,mj} h_{mj}, \tag{2.41}$$

where $\|\cdot\|_{0,ni}$ and $\|\cdot\|_{0,mj}$ denote the L^2 norm on Ω_{ni} and Ω_{mj}, $\|\cdot\|_{1,ni}$ and $\|\cdot\|_{1,mj}$ denote the H^1 norm on Ω_{ni} and Ω_{mj}, h_{ni} and h_{mj} are the diameters of the smallest balls containing the finite element Ω_{ni} and Ω_{mj} respectively, and C is a positive constant independent to h_{ni} and h_{mj}. In the next section, the approximate solution Φ_i and G_j^* are substituted into (2.33) to obtain an approximate inverse problem formulation.

2.5.2 Inverse Problem Discretization

Let $A_{ij}(\boldsymbol{r}) = G_j^*(\boldsymbol{r})\Phi_i(\boldsymbol{r})$ and $A_{ij}^*(\boldsymbol{r}) = G_j^*(\boldsymbol{r})\Phi_i(\boldsymbol{r})$, $i = 1, \ldots, N_S$ and $j = 1, \ldots, N_D$, be the finite-dimensional approximations of a_{ij} and a_{ij}^*. We substitute a_{ij} and a_{ij}^* with their finite-dimensional counterparts in the operators $\mathcal{A}$ and $\mathcal{A}^*$ defined by (2.16) and (2.26). The resulting approximate operators are denoted by tildes: $\widetilde{\mathcal{A}}$ and $\widetilde{\mathcal{A}}^*$, indicating that the finite element solutions of the forward problem are used. By so doing, we arrive at the approximate variational inverse problem formulation:

$$\widetilde{\mathcal{F}}(\psi, \widetilde{\mu}) = \widetilde{\mathcal{G}}(\psi), \quad \forall \psi \in H^1(\Omega), \tag{2.42}$$

where we use $\widetilde{\mu}$ to denote the solution of this approximate inverse problem. In (2.42), $\widetilde{\mathcal{F}}(\psi, \widetilde{\mu})$ and $\widetilde{\mathcal{G}}(\psi)$ are given respectively by

$$\begin{aligned} \widetilde{\mathcal{F}}(\psi, \widetilde{\mu}) &:= (\psi, \widetilde{\mathcal{B}}\widetilde{\mu}) + \lambda_1(\psi, \widetilde{\mu}) + \lambda_2(\nabla\psi, \nabla\widetilde{\mu}), \qquad (2.43) \\ \widetilde{\mathcal{G}}(\psi) &:= (\psi, \tilde{f}), \qquad (2.44) \end{aligned}$$

where

$$\begin{aligned} (\widetilde{\mathcal{B}}\widetilde{\mu})(\boldsymbol{r}) &:= (\widetilde{\mathcal{A}}^*\widetilde{\mathcal{A}}\widetilde{\mu})(\boldsymbol{r}) \\ &= \int_\Omega \sum_{i,j}^{N_S,N_D} A^*_{ij}(\boldsymbol{r})A_{ij}(\acute{\boldsymbol{r}})\widetilde{\mu}(\acute{\boldsymbol{r}})d\acute{\boldsymbol{r}}, \qquad (2.45) \end{aligned}$$

and

$$\begin{aligned} \tilde{f}(\boldsymbol{r}) &:= (\widetilde{\mathcal{A}}^*\boldsymbol{\Gamma})(\boldsymbol{r}) + \lambda_1\beta_1(\boldsymbol{r}) + \lambda_2\sum_{q=1}^{3}\beta_{2q}(\boldsymbol{r}) \\ &= \sum_{i,j}^{N_S,N_D} A^*_{ij}(\boldsymbol{r})\Gamma_{i,j} + \lambda_1\beta_1(\boldsymbol{r}) + \lambda_2\sum_{q=1}^{3}\beta_{2q}(\boldsymbol{r}). \qquad (2.46) \end{aligned}$$

Next, we discretize the functions ψ and $\widetilde{\mu}$. Let $V^N(\Omega) \subset H^1(\Omega)$ denote a sequence of finite-dimensional subspaces of dimension N, spanned by the first-order Lagrange basis functions $\{L_k, k = 1, \ldots, N\}$, which are associated with the set of points $\{\boldsymbol{r}_p, p = 1, \ldots, N\}$ on Ω. We replace ψ and $\widetilde{\mu}$ in (2.42) by their respective finite-dimensional counterparts $\Psi^N \in V^N(\Omega)$ and $\mu^{D,N} \in V^N(\Omega)$:

$$\Psi^N := \sum_{k=1}^{N} p_k L_k(\boldsymbol{r}), \tag{2.47}$$

$$\mu^{D,N} := \sum_{k=1}^{N} m_k L_k(\boldsymbol{r}), \tag{2.48}$$

where p_k and m_k are unknown coefficients. As it is clear that the finite-dimensional expansions are dependent on the parameter N, this dependence is hereafter suppressed. Substituting (2.47) and (2.48) into (2.42), we arrive at

$$\widetilde{\mathcal{F}}(\Psi, \mu^D) = \widetilde{\mathcal{G}}(\Psi), \quad \forall \Psi \in V(\Omega). \tag{2.49}$$

Equation (2.49) can be transformed to a matrix equation

$$\mathbf{F}_N \boldsymbol{m} = \mathbf{G}_N, \tag{2.50}$$

where $\boldsymbol{m} = [m_1, \cdots, m_N]^T$ represents the unknown coefficients in the finite expansion of (2.48), $\mathbf{F}_N$ and $\mathbf{G}_N$ are respectively the finite element matrix and the load vector resulting from the projection of (2.42) by Galerkin method.

Let $\{\Omega_t, t = 1, \ldots, N_\Delta\}$ denote the set of linear elements used to discretize (2.42), where N_Δ is the number of elements, such that $\bigcup_t^{N_\Delta} \Omega_t = \Omega$. Note that the inverse problem mesh $\{\Omega_t\}$ is independent to the meshes $\{\Omega_{ni}\}$ and $\{\Omega_{mj}\}$, which are used to discrtize the forward problem. Similar to the forward problem, a conventional error bound for error e_μ between $\widetilde{\mu}$ and μ^D on each finite element can be given by:

$$\|e_\mu\|_{0,t} \leq C \|\widetilde{\mu}\|_{1,t} h_t, \tag{2.51}$$

where $\|\cdot\|_{0,t}$ and $\|\cdot\|_{1,t}$ denote the L^2 and H^1 norms, respectively on Ω_t, h_t is the diameter of the smallest ball containing the finite element Ω_t, and C is a positive constant independent to h_t.

2.6 Discretization Error Analysis in the Noise-free Situation

In this chapter, we consider the solution of the exact inverse problem (2.33) to be the exact solution since neither the forward problem nor the inverse problem is discretized. Our objective is to examine the error in FDOT imaging due to the finite element discretization of the forward and inverse problems. Then the error analysis can be used in the design of adaptive mesh generation algorithm to reduce the discretization error in the reconstructed image. In the first section of this chapter, we present an upper bound for the error in FDOT reconstrucction due to the forward problem discretization. Therefore, the first type of error we define is $e := \mu - \widetilde{\mu}$, where $\widetilde{\mu}$ satisfies the approximate inverse problem (2.42). Note that the inverse problem is not discretized in this case. In the second section, we present an upper bound for the error in the reconstructed image resulting from the finite element discretization of the inverse problem. In this respect, we examine the error E between the solution $\widetilde{\mu}$ of approximate inverse problem (2.42) and the solution μ^D of fully-discretized inverse problem (2.49), i.e., $E = \widetilde{\mu} - \mu^D$. In this case, the error is due entirely to the discretization of the approximate inverse problem (2.42). Finally, we define the total error as the difference between μ and μ^D in terms of the two contributors:

$$\mu - \mu^D = e + E. \tag{2.52}$$

In the followings, we analyze each of these two types of errors and derive two upper bounds for the $H^1(\Omega)$ norm of these errors respectively.

2.6.1 Error due to Forward Problem Discretization

The following theorem presents an upper bound for the $H^1(\Omega)$ norm of the error e due to the forward problem discretization.

Theorem 1:

Let $\{\Omega_{ni}, n = 1, \ldots, N_\Delta^i\}$ denote the set of elements used to discretize (2.34) such that $\bigcup_n^{N_\Delta^i} \Omega_{ni} = \Omega$ and h_{ni} is the diameter of the smallest ball that contains the n^{th} element in the solution Φ_i, for all $i = 1, \ldots, N_S$. Similarly, let $\{\Omega_{mj}, m = 1, \ldots, N_\Delta^{*j}\}$ denote the set of elements used to discretize (2.35) such that $\bigcup_m^{N_\Delta^{*j}} \Omega_{mj} = \Omega$ and h_{mj} is the diameter of the smallest ball that contains the m^{th} element in the solution G_j^*, for all $j = 1, \ldots, N_D$. Then a bound for the error between the solution μ of (2.33) and the solution $\widetilde{\mu}$ of (2.42) due to the approximations of $\widetilde{\mathcal{F}}$ and $\widetilde{\mathcal{G}}$ is given by:

$$\begin{aligned} \|\mu - \widetilde{\mu}\|_1 \leq\ & \frac{C}{\min(\lambda_1, \lambda_2)} \max_{i,j} \|g_j^* \phi_i\|_0 \\ & \cdot \left(\sum_{i=1}^{N_S} \sum_{n,j}^{N_\Delta^i, N_D} (2\|g_j^* \mu\|_{0,ni} + \|g_j^*\|_{\infty,ni} \|\mu\|_0) \|\phi_i\|_{1,ni} h_{ni} \right. \\ & \left. + \sum_{j=1}^{N_D} \sum_{m,i}^{N_\Delta^{*j}, N_S} (2\|\phi_i \mu\|_{0,mj} + \|\phi_i\|_{\infty,mj} \|\mu\|_0) \|g_j^*\|_{1,mj} h_{mj} \right), \end{aligned} \tag{2.53}$$

where C is a positive constant independent of the discretization parameters h_{ni} and h_{mj}.

Proof: See Appendix D. □

The error bound in (2.53) shows the specific effect that the forward problem discretization has on the accuracy of the inverse problem solution. In this respect, for the forward problem, Theorem 1 suggests a discretization criteria for the forward problem which takes both forward and inverse problem solutions into account. First of all, the forward problem discretization includes the discretization of each solution ϕ_i and g_j^*. Second, to keep the total error bound low, the h_{ni} of the n^{th} element in solution ϕ_i has to be chosen small when $\sum_j^{N_D}(2\|g_j^*\mu\|_{0,ni} + \|g_j^*\|_{\infty,ni}\|\mu\|_0)\|\phi_i\|_{1,ni}$ is large on this element; and the h_{mj} of the m^{th} element in solution g_j^* has to be chosen small when $\sum_i^{N_S}(2\|\phi_i\mu\|_{0,mj} + \|\phi_i\|_{\infty,mj}\|\mu\|_0)\|g_j^*\|_{1,mj}$ is large on this element. Note that $\|\phi_i\|_{1,ni}$ or $\|g_j^*\|_{1,mj}$ is large on the element close to the i^{th} source or the j^{th} detector, respectively; and the values of the terms $\sum_j^{N_D}(2\|g_j^*\mu\|_{0,ni} + \|g_j^*\|_{\infty,ni}\|\mu\|_0)$ and $\sum_i^{N_S}(2\|\phi_i\mu\|_{0,mj} + \|\phi_i\|_{\infty,mj}\|\mu\|_0)$ depend on the fluorophore heterogeneity and its position with respect to the sources and detectors. Therefore, finer elements near the designated source or detector as well as near the heterogeneity may result in a lower error bound. The conventional error bounds (2.41) or (2.40) only depends on the smoothness and support of ϕ_i or g_j^* as well as the finite-dimensional space of approximating function. They only require to keep h_{ni} or h_{mj} small where $\|\phi_i\|_{1,ni}$ or $\|g_j^*\|_{1,mj}$ is large, respectively. However, lower error bound in (2.41) or (2.40) only guarantees to reduce the error in the solution of the forward problem, and may not necessarily reduce the error in the solution of inverse problem.

Furthermore, it is clear that other factors can affect this error bound. The regularization parameters scale this error bound, therefore, choosing smaller values for λ_1 and λ_2 may lead to higher discretization error. We note that, since this error bound is a sum over all sources and detectors, increasing the number of either can also have an impact on the reconstruction accuracy.

2.6.2 Error due to Inverse Problem Discretization

In the following theorem, we present an upper bound for the $H^1(\Omega)$ norm of the error E due to the inverse problem discretization.

Theorem 2:

Consider the Galerkin projection of the variational problem (2.42) on a finite-dimensional subspace $V(\Omega) \subset H^1(\Omega)$ using a set of finite elements $\{\Omega_t, t = 1, \ldots, N_\Delta\}$ whose vertices are at $\{r_p, p = 1, \ldots, N\}$, such that $\bigcup_t^{N_\Delta} \Omega_t = \Omega$, and let h_t be the diameter of the smallest ball that contains the t^{th} element. Assume that the solution $\widetilde{\mu}$ of (2.42) also satisfies $\widetilde{\mu} \in H^2(\Omega)$. Then a bound for the error E in the solution μ^D of (2.49) with respect to the solution $\widetilde{\mu}$ of (2.42) can be given by

$$\begin{aligned} \|\widetilde{\mu} - \mu^D\|_1 \quad \leq \quad & \frac{C}{\min(\lambda_1, \lambda_2)} \left(\max_{i,j} \|G_j^* \Phi_i\|_0 \sum_{t=1}^{N_\Delta} \sum_{i,j}^{N_S, N_D} \|G_j^* \Phi_i\|_{0,t} \|\widetilde{\mu}\|_{2,t} h_t^2 \right. \\ & \left. + \lambda_1 \sum_{t=1}^{N_\Delta} \|\widetilde{\mu}\|_{2,t} h_t^2 + \lambda_2 \sum_{t=1}^{N_\Delta} \|\widetilde{\mu}\|_{2,t} h_t \right), \end{aligned} \tag{2.54}$$

where C is a positive constant independent of the discretization parameters h_t.

Proof: See Appendix E. □

The error bound in (2.54) shows that the error due to the inverse problem discretization not only depends on the inverse problem solution itself, but also on the solutions of the forward problem. The first term in the bracket of (2.54) shows that $\|\widetilde{\mu}\|_{2,t}$ is scaled by the finite element solutions of the forward problem $\max_{i,j}\|G_j^*\Phi_i\|_0\|G_j^*\Phi_i\|_{0,t}$. This implies that the discretization error is dependent on the location of the heterogeneity with respect to the sources and detectors. The second and third term in the bracket suggest that keeping the mesh size small where $\|\widetilde{\mu}\|_{2,t}$ is large, can help to lower the error bound, but it also depends on the regularization parameters λ_1 and λ_2. Comparing (2.54) with the conventional error bound (2.51), (2.54) suggests a discretization criteria based on the inverse problem solution $\widetilde{\mu}$, and the forward problem solutions Φ_i and G_j^*, as well as the spatial relationship among these solutions. Because of the interdependence between the forward and inverse problems, simply keeping the mesh parameter small over regions where $\|\widetilde{\mu}\|_{2,t}$ is large, as suggested by (2.51), may not ensure a lower error bound, thereby a reduction of the error in the reconstructed image. Similar to Theorem 1, the regularization parameters λ_1, λ_2 and the number of sources and detectors also affect this error bound, therefore, may also have an impact on the reconstruction accuracy.

Combining results of Theorems 1 and 2, and rearranging the terms, both error bounds can be formulated in a single error bound. Assuming $\max_{i,j}\|G_j^*\Phi_i\|_0 \approx \max_{i,j}\|g_j^*\phi_i\|_0$ for $i = 1, \cdots, N_S$ and $j = 1, \cdots, N_D$, we have a total bound for the error due to discretization in the reconstructed image:

$$\begin{aligned}
\|\mu - \mu^D\|_1 \;\leq\; & \frac{C}{\min(\lambda_1,\lambda_2)} \max_{i,j}\|g_j^*\phi_i\|_0 \\
& \cdot \sum_{i,j}^{N_S,N_D}\left[\sum_{n=1}^{N_\Delta^i}(2\|g_j^*\mu\|_{0,ni} + \|g_j^*\|_{\infty,ni}\|\mu\|_0)\|\phi_i\|_{1,ni}h_{ni}\right. \\
& + \sum_{m=1}^{N_\Delta^{*j}}(2\|\phi_i\mu\|_{0,mj} + \|\phi_i\|_{\infty,mj}\|\mu\|_0)\|g_j^*\|_{1,mj}h_{mj} \\
& \left. + \sum_{t=1}^{N_\Delta}\|G_j^*\Phi_i\|_{0,t}\|\widetilde{\mu}\|_{2,t}h_t^2\right] \\
& + \frac{C}{\min(\lambda_1,\lambda_2)}\left(\lambda_1\sum_{t=1}^{N_\Delta}\|\widetilde{\mu}\|_{2,t}h_t^2 + \lambda_2\sum_{t=1}^{N_\Delta}\|\widetilde{\mu}\|_{2,t}h_t\right). \qquad (2.55)
\end{aligned}$$

2.6.3 Error in Iterative Linearization

The error bounds derived in the previous sections only address one iteration of linearization. If the solution of the inverse problem is iteratively updated using the method mentioned in Section 2.4, the propagation of error due to discretization needs to be considered. Let μ^k and $\mu^{D,k}$ represent the solution to (2.33) and (2.49) at the k^{th} iteration, respectively, and we know $\mu^{D,k}$ contains error with respect to μ^k. Then, in the $(k+1)^{\text{th}}$ iteration, three types of errors are introduced into the solution $\mu^{D,k+1}$: The first is due to the finite element discretization of the inverse problem. The second is due to the finite element discretization of the forward problem introducing the error $(\mathcal{B} - \widetilde{\mathcal{B}})^{k+1}$ in the operator $\widetilde{\mathcal{B}}^{k+1}$ and the error $(\widetilde{\mathcal{A}}^* - \mathcal{A}^*)^{k+1}\mathbf{\Gamma}$ in $\widetilde{\mathcal{A}}^{*,k+1}\mathbf{\Gamma}$. Finally, μ_{axf} and μ_{amf} which are related to $\mu^{D,k}$, appear as the coefficient in the forward problem (2.5)-(2.7), and the adjoint problem (2.9)-(2.10), respectively. Then, the error in the solution $\mu^{D,k}$ at each iteration will propagate and lead to an additional error in Φ_i and G_j^* at the $(k+1)^{\text{th}}$ iteration in addition to the discretization error analyzed before.

2.7 Conclusion

In this chapter, we analyzed the error in FDOT imaging due to the forward and inverse problem discretizations under the assumption that the measurements are noise-free. After giving a brief description for the forward problem of FDOT, we presented our inverse problem formulation addressing its nonlinearity by an iterative linearization method based on the Born approximation and its ill-posedness by the Tikhonov regularization in an optimization framework, respectively. Then we transformed both forward and inverse problems into variational problems, and discretized and solved them using the Galerkin projection method with first-order Lagrange basis functions. We analyzed the effect of each problem's discretization on the reconstruction accuracy. In our analysis, we first considered the impact of the forward problem discretization and provided an upper bound for the resulting error in Theorem 1. Next, we analyzed the inverse problem discretization, and obtained another upper bound for the resulting error in Theorem 2. These two new error bounds showed that the error resulting from discretization is affected by the fluorophore concentration, number of sources and detectors, and their locations with respect to the fluorophore concentration, as well as the regularization parameters. Based on these two new error bounds, we develop two new adaptive mesh generation algorithms that can be used in the adaptive FEM for FDOT imaging in the next chapter.

CHAPTER 3
ADAPTIVE MESH GENERATION FOR FDOT IN THE NOISE-FREE SITUATION[1]

3.1 Introduction

In Chapter 2, we analyzed the effect of discretization on the accuracy of FDOT reconstruction and derive two upper bounds for the error resulting from the forward and inverse problem discretizations in the noise-free situation. In this chapter, based on these two error bounds, we develop two new adaptive mesh generation algorithms which can be used in the adaptive FEM for FDOT. By selecting $\|\mu - \widetilde{\mu}\|_1$ and $\|\widetilde{\mu} - \mu^D\|_1$ as the figures of merit and reducing the respective bounds (2.53) and (2.54), we aim to minimize the error between the fully discretized solution μ^D and the exact solution μ. Clearly, we can reduce both error bounds by reducing the mesh size parameter uniformly (h_{ni} and h_{mj} in (2.53) and h_t in (2.54)), which is known as $h-$refinement in the literature. At the same time, reducing mesh size will increase the number of total discretization points, thereby leading to an increase in the size of the discretized forward and/or inverse problems.

To address this tradeoff between the accuracy and the computational complexity, we consider two adaptive h-refinement approaches: The first approach we discuss sets a *predetermined* error tolerance value for the error bound computed on each finite element, and reduces the mesh size parameter by refining the element until this *predetermined* tolerance value is reached. This approach relies on the knowledge of this error tolerance value which represents the desired level of discretization accuracy. From a practical standpoint, defining the ideal value for the error tolerance can be a complex undertaking due to the presence of unknown factors and functions such as C, g_j^*, ϕ_i, μ and $\widetilde{\mu}$ in both error bounds. The second approach, which is more easily applied in practice, limits the total number of discretization points in

---

[1] Portions of this chapter previously appeared as: M. Guven, L. Zhou, L. Reilly-Raska, and B. Yazıcı. Discretization error analysis and adaptive meshing algorithms for fluorescence diffuse optical tomography: part II. *IEEE Transactions on Medical Imaging*, 29(2):230-245, 2010.

the finite-dimensional approximations (2.36), (2.37) and (2.48). Instead of using a *predetermined* error tolerance, the average value of the error bounds on the finite elements is calculated at each iteration in the refinement process. Then the elements with computed error bounds larger than the average are refined until the limit on the discretization points number is reached. In this respect, the first approach can be viewed as to rely on *a priori* information while the second approach makes use of *a posteriori* information on the error bounds.

In the next two sections, we design the adaptive mesh generation algorithms based on the error bounds derived previously for the forward and inverse problems. First, we explore the mesh generation for the forward problem and describe the algorithm which uses a predetermined error tolerance before discussing the second approach. Next, we present the adaptive mesh generation algorithm for the inverse problem in a similar way. Unlike the adaptive meshing method suggested by the conventional error bound in finite element theory, our new algorithms address the interdependence between the solutions of the forward and inverse problems and take into account the source and detector configuration, their relative locations with respect to the fluorophore heterogeneity, and inverse problem formulation, as well as the regularization parameters. At the end of this chapter, we present a series of three-dimensional numerical simulations to illustrate the practical advantages of our new adaptive mesh generation algorithms in FDOT reconstruction.

3.2 Adaptive Mesh Generation for Forward Problem

For the forward problem discretization, our goal is to minimize $\|\mu-\widetilde{\mu}\|_1$, whose upper bound is given in (2.53). Let $\tilde{\varepsilon}_f$ be the allowable $H^1(\Omega)$ norm of the error in the reconstructed optical image due to the forward problem discretization. The bound (2.53) contains known or estimable factors multiplying two summation terms. The first term arises from the discretization of the boundary value problem (2.5)-(2.7) for each source, and the second term arises from the discretization of the boundary value problem (2.9)-(2.10) for each detector. Therefore, we need to set two distinct conditions corresponding to each of the mesh parameters h_{ni} and h_{mj} that will ensure the $H^1(\Omega)$ norm of the error is less than $\tilde{\varepsilon}_f$.

Starting with the *predetermined* allowable error norm $\tilde{\varepsilon}_f$, we define the parameter ε_f as follows:

$$\varepsilon_f := \frac{\min(\lambda_1, \lambda_2)}{C \max_{i,j} \|g_j^* \phi_i\|_0} \left(\sum_{i=1}^{N_S} N_\Delta^i + \sum_{j=1}^{N_D} N_\Delta^{*j} \right)^{-1} \tilde{\varepsilon}_f, \tag{3.1}$$

where C is the positive constant in (2.53), N_Δ^i or N_Δ^{*j} is the number of elements in the mesh used to solve Φ_i or G_j^*. Let B_{ni} and B_{mj}, $i = 1, \cdots, N_S$ and $j = 1, \cdots, N_D$, be defined as

$$B_{ni} := \frac{\varepsilon_f}{\sum_{j=1}^{N_D} (2\|g_j^* \mu\|_{0,ni} + \|g_j^*\|_{\infty,ni} \|\mu\|_0) \|\phi_i\|_{1,ni}}, \tag{3.2}$$

$$B_{mj}^* := \frac{\varepsilon_f}{\sum_{i=1}^{N_S} (2\|\phi_i \mu\|_{0,mj} + \|\phi_i\|_{\infty,mj} \|\mu\|_0) \|g_j^*\|_{1,mj}}. \tag{3.3}$$

In this case, if $h_{ni} > 0$ and $h_{mj} > 0$ are chosen as

$$\begin{aligned} h_{ni} &\leq B_{ni}, \\ h_{mj} &\leq B_{mj}^*, \end{aligned}$$

then by Theorem 1, this implies that $\|\mu - \widetilde{\mu}\|_1 \leq \tilde{\varepsilon}_f$.

For each source or detector, the algorithm is initiated with a coarse uniform mesh. Each mesh size parameter h_{ni} (or h_{mj}) in this mesh is checked against the corresponding bound B_{ni} (or B_{mj}^*). Then, each element with mesh size parameter h_{ni} (or h_{mj}) larger than the bound B_{ni} (or B_{mj}^*) is refined. With each sweep of refinement, the mesh is altered providing spatially varying resolution over the domain. The process is iterated until the mesh size parameter h_{ni} (or h_{mj}) for each element is less than the corresponding bound B_{ni} (or B_{mj}^*). In other words, the tolerance error bound $\tilde{\varepsilon}_f$ is uniformly distributed over all finite elements. Note that B_{ni} (or B_{mj}^*) is computed for a given set of meshes $\{\Delta^i, i = 1, \cdots, N_S\}$ (or $\{\Delta^{*j}, j = 1, \cdots, N_D\}$) and the associated number of elements is N_Δ^i (or N_Δ^{*j}) in each of these meshes.

We further present an alternative approach based on a different constraint. Rather than starting with a predetermined error bound as the constraint, we instead limit the total number of discretization points which determines the size of the discretized forward problem. This approach is particularly useful when an error bound can not be determined in advance and the computational resources are limited. In this approach, we neglect the factor multiplying the summation terms in (2.53), and define two error indicators for the n^{th} and m^{th} elements in the forward problem meshes associated with i^{th} source and j^{th} detector, respectively:

$$\begin{aligned} \varepsilon_f^i(n) &= \sum_{j=1}^{N_D}(2\|g_j^*\mu\|_{0,ni} + \|\mu\|_0\|g_j^*\|_{\infty,ni})\|\phi_i\|_{1,ni}h_{ni}, && (3.4) \\ \varepsilon_f^j(m) &= \sum_{i=1}^{N_S}(2\|\phi_i\mu\|_{0,mj} + \|\mu\|_0\|\phi_i\|_{\infty,mj})\|g_j^*\|_{1,mj}h_{mj}. && (3.5) \end{aligned}$$

Similar to the first approach, the adaptive mesh generation algorithm is initiated with a coarse uniform mesh. With each sweep of refinement and for each source (or detector), we compute $\varepsilon_f^i(n)$ (or $\varepsilon_f^j(m)$) on every finite element and compute the average value $\bar{\varepsilon}_f^i$ (or $\bar{\varepsilon}_f^j$) of them. Every element with $\varepsilon_f^i(n) > \bar{\varepsilon}_f^i$ (or $\varepsilon_f^j(m) > \bar{\varepsilon}_f^j$) is refined so that the new $\varepsilon_f^i(n)$ (or $\varepsilon_f^j(m)$) value for that element will decrease. In this case, the resulting mesh provides spatially varying resolution over the domain and the discretization error is more uniformly distributed over all finite elements than the last refinement iteration. The algorithm is stopped when the number of discretization points reaches the allowable limit. This approach is further described in Algorithm 1 by pseudocode.

Finally, we note that the practical implementations of both approaches require several adjustments, because B_{ni}, B_{mj}^*, $\varepsilon_f^i(n)$, and $\varepsilon_f^j(m)$ can not be computed exactly due to the unknown actual values of g_j^*, ϕ_i, μ and C. First of all, ε_f can be scaled by C to eliminate the unknown constant C. Second, B_{ni}, B_{mj}^*, $\varepsilon_f^i(n)$, and $\varepsilon_f^j(m)$ can be estimated by using the analytical solutions for g_j^*, ϕ_i on a unbounded domain (see Appendix F) and the approximation of μ based on some *a priori* information [60]. Alternatively, these functions can also be estimated using the finite-dimensional solutions G_j^*, Φ_i and μ^D from the most recent iteration of image reconstruction.

Algorithm 1 The pseudocode of the the adaptive mesh generation algorithm for the forward problem in the noise-free situation.

```
◇ Generate an initial uniform mesh:
$(\Delta^i, N^i_\Delta)$, $\Delta^i = \bigcup_{n=1}^{N^i_\Delta}\{\Delta_n\}$   $\Big((\Delta^{*j}, N^{*j}_\Delta), \Delta^{*j} = \bigcup_{m=1}^{N^{*j}_\Delta}\{\Delta_m\}\Big)$
◇ Set the maximum number of nodes $N^f_{max}$
  while Number of nodes in $\Delta^i$ $(\Delta^{*j})$ less than $N^f_{max}$
    for each element $\Delta_n \in \Delta^i$ with mesh parameter $h_{ni}$ $(\Delta_m \in \Delta^{*j}$ with $h_{mj})$
      if first linearization
        . Use analytical solutions for $\phi_i$ and $g^*_j$ and a priori information
          about $\mu$ to compute $\varepsilon^i_f(n)$ in (3.4) $(\varepsilon^j_f(m)$ in (3.5))
      else
        . Use current solution updates $\Phi_i$, $G^*_j$ and $\mu^D$ to compute $\varepsilon^i_f(n)$ in
          (3.4) $(\varepsilon^j_f(m)$ in (3.5))
    end
      . Compute $\overline{\varepsilon}^j_f$ $(\overline{\varepsilon}^i_f)$
      . Refine the elements with $\varepsilon^i_f(n) > \overline{\varepsilon}^i_f$   $(\varepsilon^j_f(m) > \overline{\varepsilon}^j_f)$
      . Update the mesh $\Delta^i$ $(\Delta^{*j})$
  end
◇ Solve for $\Phi_i$ $(G^*_j)$
```

3.3 Adaptive Mesh Generation for Inverse Problem

For the inverse problem discretization, our goal is to minimize $\|\widetilde{\mu}-\mu^D\|_1$, whose upper bound is given in (2.54). We adopt a similar approach for the inverse problem mesh generation algorithm as outlined for the forward problem previously. First, we present an algorithm for adaptive mesh generation using a predetermined error tolerance to be distributed over the whole domain. Then, we discuss the practical approach limiting the number of discretization points.

Let $\tilde{\varepsilon}_i$ be the *predetermined* allowable upper bound of the error in the reconstructed image due to the Galerkin projection of (2.42). We aim to distribute this error bound evenly over all elements in the domain. Let

$$\varepsilon_i := \frac{\min(\lambda_1, \lambda_2)}{CN_\Delta}\tilde{\varepsilon}_i. \tag{3.6}$$

and

$$
\begin{aligned}
a(t) &:= \max_{i,j} \|G_j^* \Phi_i\|_0 \sum_{i,j}^{N_S, N_D} \|G_j^* \Phi_i\|_{0,t} \|\widetilde{\mu}\|_{2,t} + \lambda_1 \|\widetilde{\mu}\|_{2,t}, && (3.7)\\
b(t) &:= \lambda_2 \|\widetilde{\mu}\|_{2,t}, && (3.8)
\end{aligned}
$$

such that Theorem 2 can be rewritten as follows:

$$\|\widetilde{\mu} - \mu^D\|_1 \le \frac{C}{\min(\lambda_1, \lambda_2)} \sum_{t=1}^{N_\Delta} \left[a(t) h_t^2 + b(t) h_t \right]. \tag{3.9}$$

Next, we define the parameter B_t as follows:

$$B_t := \frac{-b(t) + \sqrt{b(t)^2 + 4a(t)\varepsilon_i}}{2a(t)}. \tag{3.10}$$

Then for $0 < h_t \le B_t$, by (3.9) and (3.6), we have

$$\|\widetilde{\mu} - \mu^D\|_1 \le \frac{C N_\Delta}{\min(\lambda_1, \lambda_2)} \varepsilon_i \le \tilde{\varepsilon}_i.$$

We note that the adaptive mesh generation using this approach for the inverse problem remains the same as the one we outlined for the forward problem: starting from a coarse uniform mesh, every element in the mesh is refined until its size parameter satisfies $h_t \le B_t$.

The second more practical approach limits the number of discretization points in the adaptive mesh rather than relying on a predetermined error tolerance. As in the previous section, this algorithm uses an error indicator computed after each sweep of refinement on every finite element:

$$\varepsilon_i(t) := a(t) h_t^2 + b(t) h_t, \tag{3.11}$$

where $a(t)$ and $b(t)$ are defined in (3.7) and (3.8), respectively. Given an initial coarse uniform mesh, we compute ε_i for each element with parameter h_t and compute the average value $\overline{\varepsilon}_i$ of them. Next, we refine those elements with $\varepsilon_i(t) > \overline{\varepsilon}_i$, so that ε_i computed on the new element becomes smaller. The algorithm has to be stopped before the total number of discretization points exceeds the allowable limit. Computing B_t or ε_i requires knowledge of $\widetilde{\mu}$. In this respect, we can either estimate $\widetilde{\mu}$ using *a priori* information, or use the most recent updates of the inverse problem solutions μ^D as an approximation. The pseudocode for this algorithm is outlined in Algorithm 2.

Algorithm 2 The pseudocode of the the adaptive mesh generation algorithm for the inverse problem in the noise-free situation.

```
◇ Generate an initial uniform mesh: $(\Delta, N_\Delta)$, $\Delta = \bigcup_{t=1}^{N_\Delta}\{\Delta_t\}$
◇ Set the maximum number of nodes $N^i_{max}$
  while Number of nodes $N$ less than $N^i_{max}$
    for each element $\Delta_t \in \Delta$ with mesh size parameter $h_t$
      if first linearization
        . Use current solution updates $\Phi_i$ and $G^*_j$ and a priori information
          about $\widetilde{\mu}$ to compute $\varepsilon_i(t)$ in (3.11)
      else
        . Use current solution updates $\Phi_i$ and $G^*_j$ and $\mu^D$ to compute $\varepsilon_i(t)$ in
          (3.11)
    end
      . Compute $\overline{\varepsilon}_i$
      . Refine the elements with $\varepsilon_i(t) > \overline{\varepsilon}_i$
      . Update the mesh $\Delta$
  end
◇ Solve for $\mu^D$.
```

3.4 A Comparison to the Conventional Techniques

There are various *a priori* error bounds developed in particular to estimate discretization error in the numerical solutions of PDEs (see [67] for a survey on error estimation procedures). These error bounds are in general adopted for adaptive mesh generation in inverse problems. One commonly used is the interpolation error bound [79, 94, 96], which depends on the smoothness of the solution, the mesh size, and the basis function (first-order Lagrangian basis function in our case) used for approximation:

$$\|\widetilde{\mu} - \widetilde{\mu}_I\|_{1,t} \leq C\|\widetilde{\mu}\|_{2,t} h_t, \tag{3.12}$$

where $\widetilde{\mu}_I$ is the interpolant of $\widetilde{\mu}$, and C is a constant similar as those in (2.53) and (2.54). Adoption of such error bounds for adaptive mesh generation in imaging problems overlooks the intrinsic mechanisms in the problem, such as the underlying forward problem, the source-detector configuration, the location of the fluorophore concentration with respect to the sources and detectors, and the regularization parameters in the inverse problem formulation. On the other hand, the error bound (2.54) in Theorem 2, takes into account these factors in addition to the smoothness of the inverse problem solution, the mesh size and the approximation basis function. In the following discussion, we compare the *a priori* interpolation error bound with our error bound (2.54) when they are used to assess the potential error in the reconstructed optical images, resulting from the discretization of the inverse FDOT problem.

Consider the discretization of the inverse problem (2.42) on a finite-dimensional subspace $V(\Omega) \subset H^1(\Omega)$ spanned by a set of linear finite elements $\{\Omega_t, t = 1, ..., N_\Delta\}$ such that $\bigcup_t^{N_\Delta} \Omega_t = \Omega$. The vertices of $\{\Omega_t, t = 1, ..., N_\Delta\}$ are at $\{r_p, p = 1, ..., N\}$ and h_t is the diameter of the smallest ball that contains the t^{th} element Ω_t. Further assume that, the following constraint is satisfied for the *a priori* interpolation error bound on the linear finite element Ω_t of the mesh for $t = 1, ..., N_\Delta$:

$$C\|\widetilde{\mu}\|_{2,t} h_t \leq \epsilon, \tag{3.13}$$

so that

$$\|\widetilde{\mu} - \mu^D\|_1 \leq C \sum_{t=1}^{N_\Delta} \|\widetilde{\mu}\|_{2,t} h_t \leq N_\Delta \epsilon. \tag{3.14}$$

In other words, if the mesh used in discretizing the inverse problem (2.42) is generated with every element subject to (3.13) which is based on the conventional error bound, then the upper bound on the $H^1(\Omega)$ norm of the error $\widetilde{\mu} - \mu^D$ is expected not to exceed $N_\Delta \epsilon$.

In the following, we show how a constraint on the conventional error bound as given in (3.13) leads to a bound on $\|\widetilde{\mu} - \mu^D\|_1$ that is higher than expected, when the parameters specific to optical imaging (such as the forward problem solution, the source-detector configuration, the inverse problem formulation) as noted in Theorem 2 are taken into account. We note that such a higher error bound may indicate a relatively higher error in the reconstructed optical image μ^D.

Substituting the constraint (3.13) into the bound (2.54) in Theorem 2, we obtain a new upper bound Θ for $\|\widetilde{\mu} - \mu^D\|_1$ and a lower bound on Θ can be given as (see Appendix G):

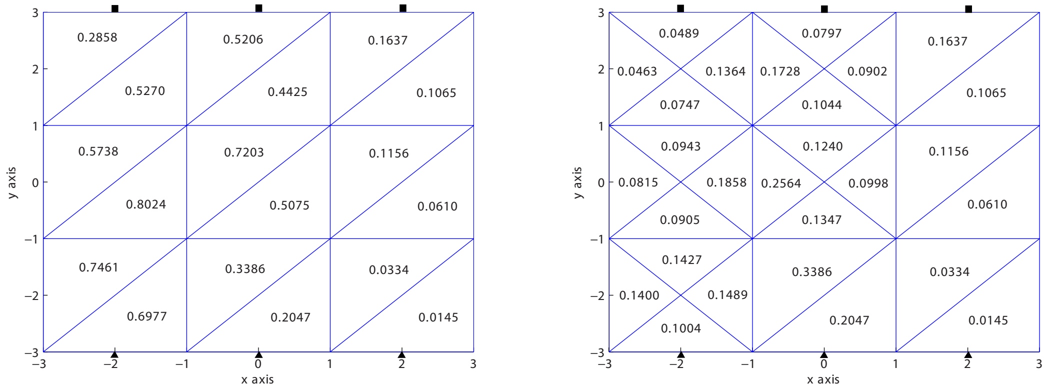
$$\begin{aligned} \Theta &\geq N_\Delta \epsilon \left[1 + \min_t h_t + \frac{\max_{i,j} \|G_j^* \Phi_i\|_0}{\min(\lambda_1, \lambda_2)} \min_t \sum_{i,j}^{N_S, N_D} \|G_j^* \Phi_i\|_{0,t} h_t \right] \\ &\geq N_\Delta \epsilon. \end{aligned} \tag{3.15}$$

Depending on the choices of λ_1 and λ_2, the last term in brackets can dominate the lower bound on Θ, or can be neglected. In either case, the lower bound in (3.15) implies that the mesh generation based on the conventional error bound indicates an increase in the error bound with respect to the expected bound of $N_\Delta \epsilon$, which may result in error in the reconstructed optical image higher than intended for.

3.5 Adaptive Mesh Generation for A Toy Problem

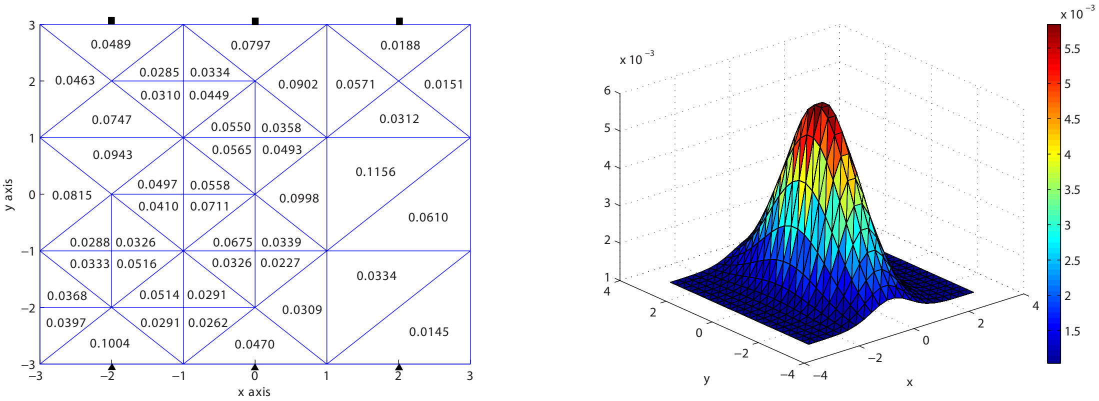
In this section, we demonstrate the mesh refinement process of our algorithms with a simple example. We consider a two-dimensional bounded domain discretized by an initial uniform mesh with 16 nodes and 18 elements as shown in Figure 3.1(a). The sources and detectors are indicated by the solid triangles and squares, respectively. Using the adaptive mesh generation algorithm described in Algorithm 1, we want to generate an adaptive mesh with triangular finite elements to solve the boundary value problem for a point source located at $(-2.5, -3)$. We constrain the number of nodes in the final mesh not to exceed 32, which represents the computational resources available at hand.

We assume that we do not have the finite element solutions of the forward problem: Φ_i, $i = 1, \cdots, N_S$, and G_j^*, $j = 1, \cdots, N_D$. Thus, we use the analytical solution of the diffusion equation on an unbounded domain to approximate ϕ_i and



(a) The initial uniform mesh with 16 nodes and 18 elements.

(b) The adaptive mesh with 21 nodes and 28 elements after first refinement.



(c) The adaptive mesh with 30 nodes and 45 elements after second refinement.

(d) The *a priori* image model for μ.

Figure 3.1: The adaptive mesh generation for a point source located at $(-2.5, -3)$ with *a priori* fluorophore concentration image shown in the last figure. The error indicator values computed for each element based on (3.4) are shown inside each element.

g_j^*, and use an *a priori* image model given by

$$\mu(x, y) = \mu_0 e^{-x^2-(\frac{y}{2})^2} + 0.001, \tag{3.16}$$

where $\mu_0 = 0.005\text{cm}^{-1}$, to approximate μ (shown in Figure 3.1(d)). Note that this *a priori* image model is a two-dimensional Gaussian function centered at $(0, 0)$. We compute the error indicators (3.4) on each of the 18 elements in the initial uniform mesh. In Figure 3.1(a), the computed error indicator values are shown inside the corresponding finite elements. The elements with error indicator computed higher than the average are selected for refinement. Using Rivara's algorithm [97], the adaptive refinement results in a mesh with 21 nodes and 28 elements as shown in Figure 3.1(b). We notice that the refinement occurs around the source at $(-2.5, -3)$ and around the detectors which are relatively close to that source. The error indicators are re-computed on the refined mesh, which are shown inside the elements in Figure 3.1(b). Note that the newly computed values are lower as compared to the previously computed ones where the elements are refined. After the next refinement, we see that the mesh is refined around the source, the closest detectors, and the fluorophore heterogeneity (see Figure 3.1(c)). At this point, the number of nodes reaches to 30, and the refinement is terminated since the number of nodes exceeds the limit if one more refinement is performed.

3.6 Computational Complexity

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 algorithm for the forward problem, which is followed by the algorithm for the inverse problem.

Using first-order Lagrange basis functions and an analytical (exact) integration on each finite element, we assume the number of multiplications required to compute the L^2 or H^1 norm of a finite-dimensional function on each element is of $\mathcal{O}(1)$ complexity [60]. Then the computational complexity of computing (3.4) is $\mathcal{O}(N_D)$ for each element. For all elements in the mesh, the computational complexity becomes $\mathcal{O}(N_\Delta^i N_D)$. Similarly, one can obtain the computational complexity of $\mathcal{O}(N_\Delta^{*j} N_S)$ for computing the error indicator (3.5). It is possible to reduce the computational complexities by making the following approximations on the error indicators (3.4) and (3.5):

$$\varepsilon_f^i(n) \approx \left(2 \left\| \sum_{j=1}^{N_D} g_j^* \mu \right\|_{0,ni} + \|\mu\|_0 \sum_{j=1}^{N_D} \|g_j^*\|_{\infty,ni} \right) \|\phi_i\|_{1,ni} h_{ni}, \tag{3.17}$$

$$\varepsilon_f^j(m) \approx \left(2 \left\| \sum_{i=1}^{N_S} \phi_i \mu \right\|_{0,mj} + \|\mu\|_0 \sum_{i=1}^{N_S} \|\phi_i\|_{\infty,mj} \right) \|g_j^*\|_{1,mj} h_{mj}. \tag{3.18}$$

Then, the computational complexity for each mesh refinement sweep using (3.17) or (3.18) reduces to $\mathcal{O}(N_\Delta^i)$ or $\mathcal{O}(N_\Delta^{*j})$, respectively.

Next, we discuss the computational complexity of our algorithm for the inverse problem. Based on the same arguments, the computational complexity of computing (3.11) is $\mathcal{O}(N_S N_D)$ for each element. For all elements in the mesh, the computational complexity becomes $\mathcal{O}(N_\Delta N_S N_D)$. Similarly, it is possible to reduce the complexity by making the following approximation on the error indicator (3.11):

$$\varepsilon_i(t) \approx \left(\max_{i,j} \|G_j^* \Phi_i\|_0 \left\| \sum_{i,j}^{N_S,N_D} G_j^* \Phi_i \right\|_{0,t} \|\mu\|_{2,t} + \lambda_1 \|\mu\|_{2,t} \right) h_t^2 + \lambda_2 \|\mu\|_{2,t} h_t. \tag{3.19}$$

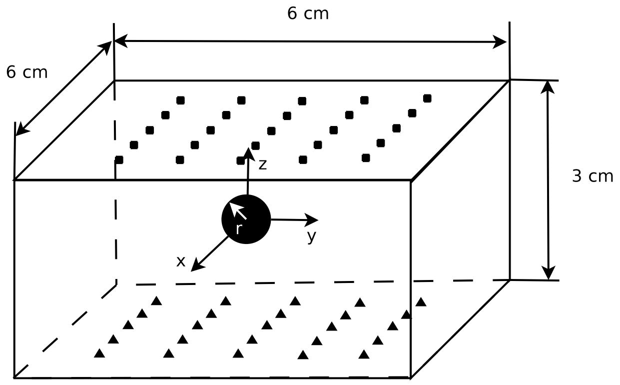
Then the computational complexity for each sweep of the mesh refinement using (3.19) reduces to $\mathcal{O}(N_\Delta)$. Note that, for conventional method, the computational complexity for one sweep of the mesh refinement is always $\mathcal{O}(N_\Delta^i)$, $\mathcal{O}(N_\Delta^{*j})$ or $\mathcal{O}(N_\Delta)$. Although it generally requires less computational resources to compute the conventional error bounds, the resulting meshes may not be as effective as the adaptive meshes generated based on our error bounds in reducing the discretization error in the reconstructed optical images.

3.7 Numerical Simulation

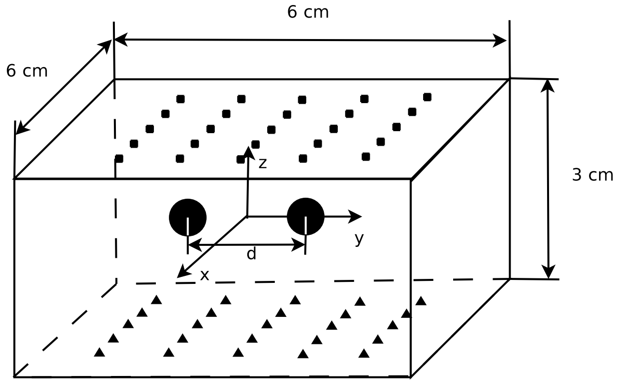
To demonstrate the improvements in the reconstructed image quality by using the adaptive FEM with our new adaptive mesh generation algorithms, we performed a series of numerical simulation studies. Specifically, we evaluated the quality of the reconstructed images in terms of reconstruction accuracy, small target detectability, and resolution as well as convergence, and compared our results with those of the uniform and conventional adaptive meshing schemes.

3.7.1 Simulation Setup

We considered the three-dimensional bounded domain shown in Figure 3.2(a) in our numerical experiments. The spherical heterogeneity with radius r in the figure denotes the concentration of the fluorophore with constant absorption coefficient



(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 sphere is 2 mm.

Figure 3.2: The simulation setups used in Simulation Studies 1 and 2. The squares and triangles denote the detectors and sources, respectively.

μ_{axf} and quantum efficiency $\eta = 0.05$, embedded in an optically homogeneous background with $\mu_{axe} = \mu_{ame} = 0.05$ cm^{-1} at both excitation and emission wavelengths. We set the diffusion coefficient $D(\boldsymbol{r}) = 0.0410$ for $\boldsymbol{r} \in \Omega \cup \partial\Omega$, the refractive index mismatch parameter $\rho = 3$ for the boundary, and placed 25 sources and 25 detectors evenly at the bottom and top surfaces of the domain. Using the parameters above, we simulated the fluorescence data by solving the coupled diffusion equations (2.5) and (2.6) with their corresponding boundary conditions (2.7) and (2.8) on a fine uniform grid with 81×81×41 nodes.

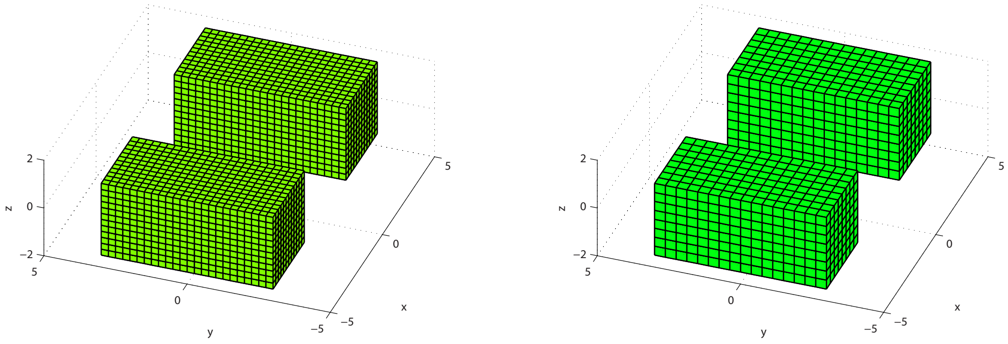
In Simulation Study 1, we considered the geometry shown in Figure 3.2(a). We set $\mu_{axf} = 0.015$ cm^{-1} and simulated 5 different data sets corresponding to 5 different values of the radius r: 1, 2, 3, 4, and 5 mm, respectively. By varying the size of the fluorophore heterogeneity, we demonstrated how our adaptive mesh generation algorithms can improve the reconstruction accuracy as well as the small target detectability of the imaging system in comparison with uniform and conventional adaptive meshing schemes.

In Simulation Study 2, we considered the geometry shown in Figure 3.2(b), and set $\mu_{axf} = 0.015$ cm^{-1}, $r = 2$ mm for both fluorophore heterogeneities inside the imaging domain. We assumed that the center of the domain is positioned at $(0, 0, 0)$ and placed two heterogeneities on the y-axis with equal distances to the origin. We simulated 5 different data sets corresponding to 5 different distances, d, between the two heterogeneities, i.e., $d = 0.5, 0.75, 1.0, 1.25, 1.5$ cm. By changing the distance between two fluorophore heterogeneities, we compared the accuracy as well as the resolution of the images reconstructed by using our adaptive mesh generation algorithms as well as the uniform and conventional adaptive meshing schemes.

In the image reconstruction, we chose the regularization parameters in our inverse problem formulation as small as possible, yet large enough to enable a robust image reconstruction. In this respect, the appropriate values for the regularization parameters were empirically selected as $\lambda_1 = 1 \times 10^{-8}$ and $\lambda_2 = 1 \times 10^{-8}$ for both Simulation Studies 1 and 2. We performed our simulation studies using deal.II FEM C++ library [98], and used hexahedral finite elements with trilinear Lagrange basis functions to discretize both forward and inverse problems. Note that we used the Gaussian quadrature method to evaluate the integrals in the variational problems (2.34), (2.35) and (2.49) [99]. While solving the forward (or inverse) problem, we evaluated the value of the inverse (forward) problem solution at the Gaussian quadrature points associated with the forward (inverse) problem mesh.

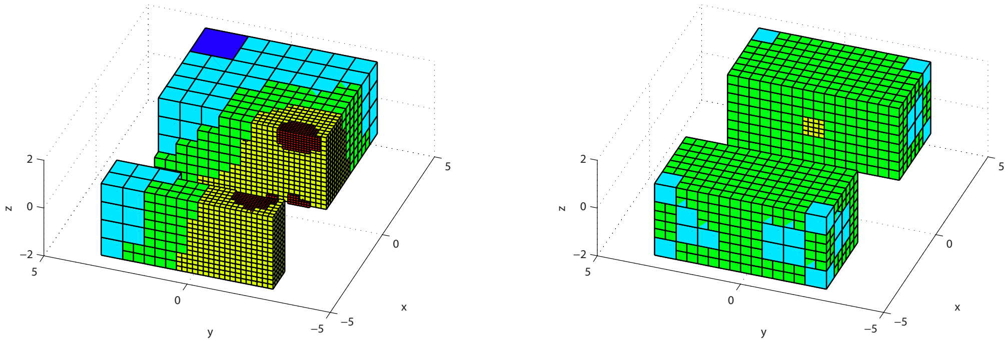
3.7.2 Mesh Generation

We used three different types of coarse meshes: Uniform mesh, the adaptive mesh generated based on the conventional error bounds (2.40), (2.41), and (3.13), and the adaptive mesh generated by our algorithms based on Theorems 1 and 2, to discretize the forward and inverse problems. For the forward problem, the number of nodes for the coarse mesh ranged from 7,000 to 9,000; and for the inverse problem, it



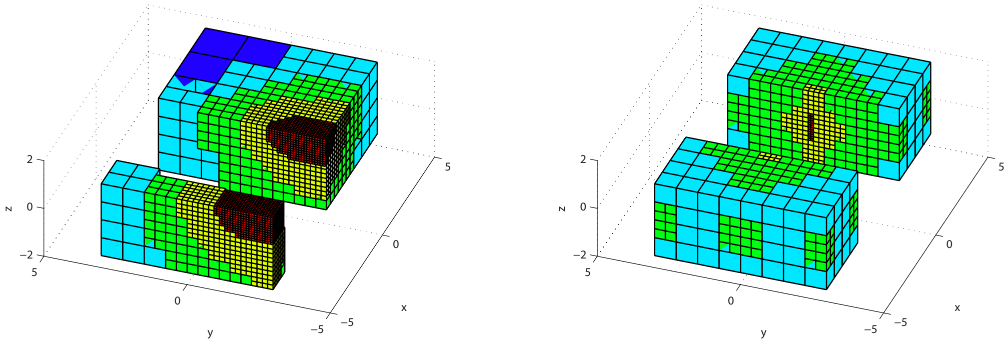
(a) The coarse uniform mesh with 8,125 nodes used to discretize the forward problem.

(b) The coarse uniform mesh with 2,601 nodes used to discretize the inverse problem.



(c) The adaptive mesh with 8,141 nodes generated for the forward problem for the detector located at (-2.0,-2.0,1.5) in Simulation Study 1, Case 5, with $\mu_{af} = 0.015$ cm^{-1}, $r = 5$ mm.

(d) The adaptive mesh with 2,583 nodes generated for the inverse problem in Simulation Study 1, Case 5, with $\mu_{af} = 0.015$ cm^{-1}, $r = 5$ mm.



(e) The conventional adaptive mesh with 8,756 nodes generated for the forward problem for the detector located at (-2.0,-2.0,1.5) in Simulation Study 1, Case 5, with $\mu_{af} = 0.015$ cm^{-1}, $r = 5$ mm.

(f) The conventional adaptive mesh with 2,691 nodes generated for the inverse problem in Simulation Study 1, Case 5, with $\mu_{af} = 0.015$ cm^{-1}, $r = 5$ mm.

Figure 3.3: Examples of the meshes used in Simulation Studies 1 and 2. The mesh is cut through to show the mesh structure inside.

ranged from 2,000 to 3,000. For performance evaluation, we considered five different image reconstruction scenarios corresponding to five different meshing schemes:

1. We used the coarse uniform mesh shown in Figure 3.3(a) to discretize the forward problem and the coarse uniform mesh shown in Figure 3.3(b) to discretize the inverse problem. We denote the resulting image by μ^D_{UU}.
2. We used the adaptive meshes generated by Algorithm 1 (see Figure 3.3(c)) to discretize the forward problem and the coarse uniform mesh shown in Figure 3.3(b) to discretize the inverse problem. We denote the resulting image by μ^D_{AU}.
3. We used the adaptive meshes generated by Algorithm 1 (see Figure 3.3(c)) to discretize the forward problem and the adaptive mesh generated by Algorithm 2 (see Figure 3.3(d)) to discretize the inverse problem. We denote the resulting image by μ^D_{AA}.
4. We used the adaptive meshes generated based on the conventional error bounds given in (2.41) and (2.40) (see Figure 3.3(e)) to discretized the forward problem and the coarse uniform mesh shown in Figure 3.3(b) to discretize the inverse problem. We denote the resulting image by μ^D_{CAU}.
5. We used the adaptive meshes generated based on the conventional error bounds given in (2.41), (2.40) and (3.13) to discretized the forward and inverse problems (see Figure 3.3(f)). We denote the resulting image by μ^D_{CAA}.

The coarse uniform mesh used for solving the forward problem shown in Figure 3.3(a) has 25×25×13 nodes and the uniform mesh used for solving the inverse problem shown in Figure 3.3(b) has 17×17×9 nodes. For adaptive meshes, we first generated the meshes for the forward and inverse problems using Algorithms 1 and 2 described in Sections 3.2 and 3.3. Then, we generated the conventional adaptive meshes for the forward and inverse problems based on the error bounds (2.41), (2.40) and (3.13).

Figures 3.3(c) and 3.3(e) show examples of our adaptive mesh and the conventional adaptive mesh for the detector located at $(-2.0, -2.0, 1.5)$. We observe that our algorithm refined the mesh around the designated detector at $(-2.0, -2.0, 1.5)$, and fluorophore heterogeneity, as well as some sources close to this detector on the opposite side of the domain. However, the conventional method only refined the mesh based on the forward problem solution associated with that detector. Figures 3.3(d) and 3.3(f) show examples of our adaptive mesh and the conventional adaptive mesh for the inverse problem when $\mu_{axf} = 0.015$ cm^{-1} and $r = 1$ mm. We observe that the conventional method refined the mesh only based on the fluorophore concentration. Our algorithm, on the other hand, performed the refinement based on the fluorophore concentration as well as the position of the fluorophore heterogeneity with respect to the sources and detectors.

In the following sections, we discuss the simulation results and compare our adaptive mesh generation algorithms with the uniform and conventional meshing schemes in terms of reconstruction accuracy, small target detectability, and resolution, as well as convergence rate of the discretized inverse problem solution to the exact inverse problem solution.

3.7.3 Simulation Results

3.7.3.1 Reconstruction Accuracy

To evaluate the reconstruction accuracy, we used a fine uniform mesh with $61\times61\times31$ nodes to discretize the forward problem (2.5), (2.8) and the inverse problem (2.33). We assumed that the error due to discretization in the resulting image, denoted by μ, is negligible with respect to the images reconstructed using the coarse meshes, and we used this image as a baseline for comparison. For both Simulation Studies 1 and 2, we summarized the $H^1(\Omega)$ norm of the error in the reconstructed images with respect to the baseline image in Tables 3.1 and 3.2. Additionally, we tabulated the percentage of error as compared to the error in the image reconstructed by using the coarse uniform meshes for the forward and inverse problems. In Tables 3.1 and 3.2, the upper row in each quantity is the absolute error value, and the lower row in each quantity is the percentage of error.

The results show that the error in the images reconstructed by using the uniform meshes for both forward and inverse problems is significantly reduced when the

Table 3.1: $H^1(\Omega)$ norms of the errors due to discretization in the images reconstructed by using different meshes in Simulation Study 1.

	Case 1	Case 2	Case 3	Case 4	Case 5
r (mm)	1	2	3	4	5
$\|\mu - \mu^D_{UU}\|_1$ ($\times 10^{-5}$)	2.315 100%	3.906 100%	5.627 100%	6.913 100%	8.632 100%
$\|\mu - \mu^D_{AU}\|_1$ ($\times 10^{-5}$)	1.417 61.21%	2.340 59.91%	3.322 59.03%	3.978 57.54%	4.810 55.72%
$\|\mu - \mu^D_{AA}\|_1$ ($\times 10^{-5}$)	0.930 40.17%	1.826 46.75%	2.852 50.69%	3.611 52.24%	4.483 51.93%
$\|\mu - \mu^D_{CAU}\|_1$ ($\times 10^{-5}$)	2.635 113.81%	4.184 107.12%	5.550 98.63%	6.259 90.54%	7.656 88.69%
$\|\mu - \mu^D_{CAA}\|_1$ ($\times 10^{-5}$)	2.525 109.09%	4.076 104.34%	5.390 95.78%	5.997 86.75%	7.338 85.01%

adaptive meshes generated by our algorithms are used. We note that the adaptive meshes generated by our algorithms reduces the error up to around 60% when the radius of heterogeneity is close to 1 mm. Although the error increases as the size of heterogeneity increases, the total error reduction provided by our algorithms is roughly around 50%. On the other hand, the conventional adaptive meshes even increases the error by roughly 5% when the size of the heterogeneity is small, and provides only about 15% error reduction when the radius of the heterogeneity increases to 5 mm.

3.7.3.2 Detectability

The reconstructed image of a point-like heterogeneity can be viewed as a point spread function of the reconstruction algorithm. To evaluate the detectability of small heterogeneities in reconstructed images using different mesh types, we consider the "peak-to-sidelobe ratio" and the Full-Width-at-Half-Maximum (FWHM) of the point spread function.

To capture peak-to-sidelobe ratio as a figure of merit, we define the Signal-to-Background-Ratio (SBR) as the ratio of the magnitude of the image at the target location to the magnitude of surrounding background volume. More specifically, we

Table 3.2: $H^1(\Omega)$ norms of the errors due to discretization in the images reconstructed by using different meshes in Simulation Study 2.

	Case 1	Case 2	Case 3	Case 4	Case 5
d (cm)	0.5	0.75	1.0	1.25	1.5
$\|\mu - \mu^D_{UU}\|_1$ ($\times 10^{-5}$)	4.392 100%	4.655 100%	4.512 100%	4.670 100%	4.733 100%
$\|\mu - \mu^D_{AU}\|_1$ ($\times 10^{-5}$)	2.647 60.27%	2.768 59.46%	2.756 61.09%	2.840 60.82%	2.910 61.48%
$\|\mu - \mu^D_{AA}\|_1$ ($\times 10^{-5}$)	1.990 45.31%	2.092 44.95%	2.077 46.03%	2.142 45.87%	2.147 45.36%
$\|\mu - \mu^D_{CAU}\|_1$ ($\times 10^{-5}$)	4.746 108.06%	4.979 106.95%	4.804 106.48%	4.988 106.80%	5.071 107.17%
$\|\mu - \mu^D_{CAA}\|_1$ ($\times 10^{-5}$)	4.577 104.22%	4.821 103.57%	4.660 103.29%	4.866 104.19%	4.947 104.53%

define

$$\mathrm{SBR}[\mu^D] = 20\log_{10}\frac{(\|\mu^D\|^2_{L^2(\Omega_F)}/V_{\Omega_F})}{(\|\mu^D\|^2_{L^2(\Omega_B)}/V_{\Omega_B})}, \tag{3.20}$$

where Ω_F and Ω_B denote the target and background regions, V_{Ω_F} and V_{Ω_B} denote the corresponding volumes.

We computed the FWHM for the cross-sections of the reconstructed images along the x- and y-axis since the FWHM for the cross-section along the z-axis is expected to be poor due to the geometry of the source-detector distribution. For each cross-section, we computed the distance between the two points corresponding to the half of the maximum fluorophore concentration. The overall FWHM of the image was determined by the average of the FWHMs along the x- and y-axis.

In Simulation Study 1, we summarized the SBR and FWHM of the reconstructed images in Tables 3.3 and 3.4. The results show that the images reconstructed using our adaptive mesh generation algorithms have higher SBR and lower FWHM than the those reconstructed using the uniform and conventional meshing schemes.

Figure 3.4 shows the cross-sections of the reconstructed images on $z = 0$ plane for the fluorophore heterogeneity with 1 mm radius. The cross-section of the baseline image, μ, which was used to compute the error values in Table 3.1, is shown in Figure 3.4(a). Figures 3.4(b), 3.4(c) and 3.4(e) show the cross-sections of

Table 3.3: SBR of the images reconstructed by using different meshes in Simulation Study 1.

	Case 1	Case 2	Case 3	Case 4	Case 5
r (mm)	1	2	3	4	5
μ (dB)	31.18	37.27	42.30	47.06	50.07
μ^D_{UU} (dB)	17.32	23.12	28.53	32.97	36.84
μ^D_{AU} (dB)	25.70	31.06	37.76	42.15	45.13
μ^D_{AA} (dB)	27.05	32.27	39.64	44.06	46.41
μ^D_{CAU} (dB)	15.34	22.17	29.91	35.09	40.55
μ^D_{CAA} (dB)	16.89	23.75	30.89	35.54	41.95

Table 3.4: FWHM of the images reconstructed by using different meshes in Simulation Study 1.

	Case 1	Case 2	Case 3	Case 4	Case 5
r (mm)	1	2	3	4	5
μ (mm)	6.04	8.75	10.98	12.30	13.71
μ^D_{UU} (mm)	7.82	9.32	12.84	13.16	14.72
μ^D_{AU} (mm)	6.66	8.81	11.34	12.79	14.04
μ^D_{AA} (mm)	6.54	8.77	11.27	12.75	13.95
μ^D_{CAU} (mm)	8.46	11.29	11.79	13.12	14.63
μ^D_{CAA} (mm)	8.56	11.25	11.74	13.05	13.99

the images μ^D_{UU}, μ^D_{AU}, and μ^D_{CAU}, reconstructed when the coarse uniform mesh was used in solving the inverse problem. We see that the concentration of fluorophore heterogeneity was reconstructed in a pyramid shape in all images. Since in this case the actual size of the fluorophore heterogeneity is smaller than the element size (3.75 mm for each edge), the coarse uniform mesh was not able to resolve the actual shape of the fluorophore heterogeneity. Figures 3.4(d) and 3.4(f) show the images of μ^D_{AA} and μ^D_{CAA}, reconstructed using the adaptive meshes for both forward and inverse problems. Since both our adaptive mesh generation algorithm and the conventional method adaptively refined the inverse problem mesh around the fluorophore heterogeneity, its shape was better resolved than that in μ^D_{UU}, μ^D_{AU}, and μ^D_{CAU}. Furthermore, comparing all the images reconstructed using the coarse meshes, we observe that the cross-sections of μ^D_{AU} and μ^D_{AA} have the largest SBR and the smallest FWHM as compared to the other images, which is consistent with the results in Table 3.3.

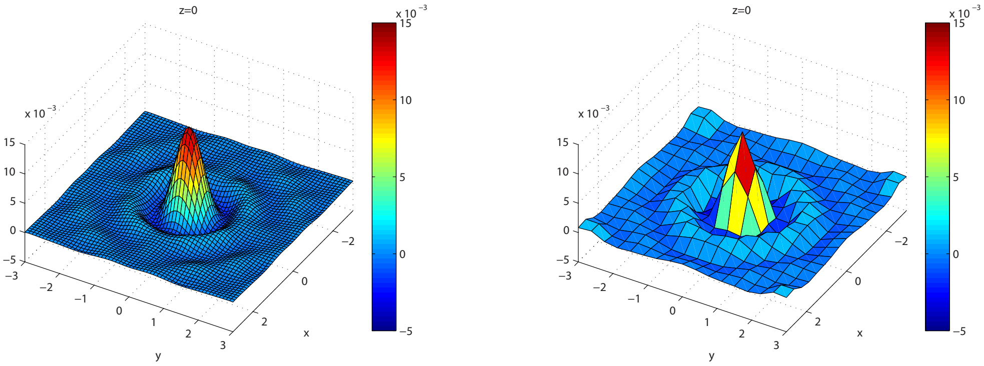
Figure 3.5(a) shows the reconstructed images of μ, μ^D_{UU}, μ^D_{AA} and μ^D_{CAA} along the y-axis. The solid line in Figure 3.5(a) shows the baseline image μ which is assumed to have no or negligible error. We observe that the image μ^D_{AA} is the closest one to μ among all three reconstructed images. In particular, we observe that this image has higher response at the center of the fluorophore heterogeneity and lower sidelobe magnitude at the background region, as compared to those reconstructed by using the conventional adaptive meshes and uniform meshes. The results in this part of the simulation study indicate that our adaptive mesh generation algorithms can effectively improve the detectability of small targets as compared to the other meshing schemes.

3.7.3.3 Resolution

To quantify the resolvability of two closely spaced heterogeneities in the reconstructed images, we define a figure of merit which we referred to as Peak-to-Valley-Ratio (PVR). PVR is the ratio between the average reconstructed concentration of the fluorophore at the two peak values corresponding to the true fluorophore locations and that at the lowest value between the two peaks. We summarized the PVR of the reconstructed images in Table 3.5.

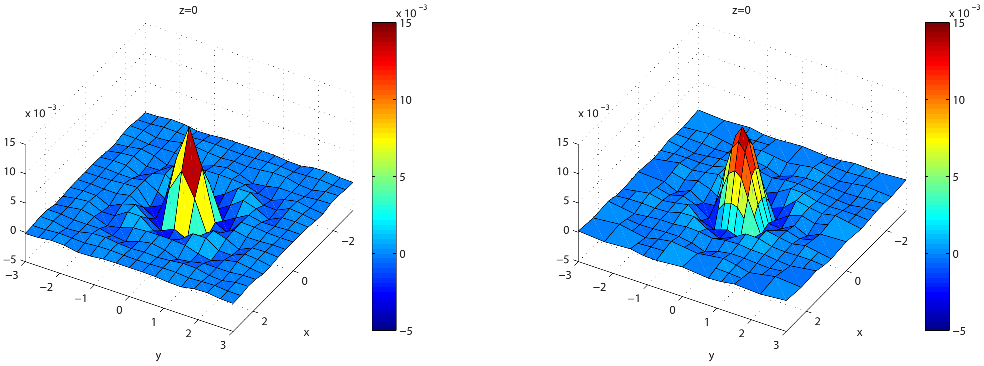
We observe that the two fluorophore heterogeneities are not distinguishable even in the baseline images, when they are closer than 1 cm apart. We note that this is due to the physical limitations of the photon propagation based on the parameters used in our simulations. We further note that, when the conventional adaptive meshes are used, the two fluorophore concentrations become distinguishable in the reconstructed images only if the distance between them is at least 1.25 cm. On the other hand, this distance is 1 cm for images reconstructed using our adaptive meshes and the uniform meshes. Among the images, in which the two fluorophore heterogeneities are distinguishable, the quantitative results show that those reconstructed using our adaptive meshes have higher PVR than those reconstructed by using the uniform and conventional adaptive meshes.

In Figure 3.6, we present the cross-sections of the reconstructed images on



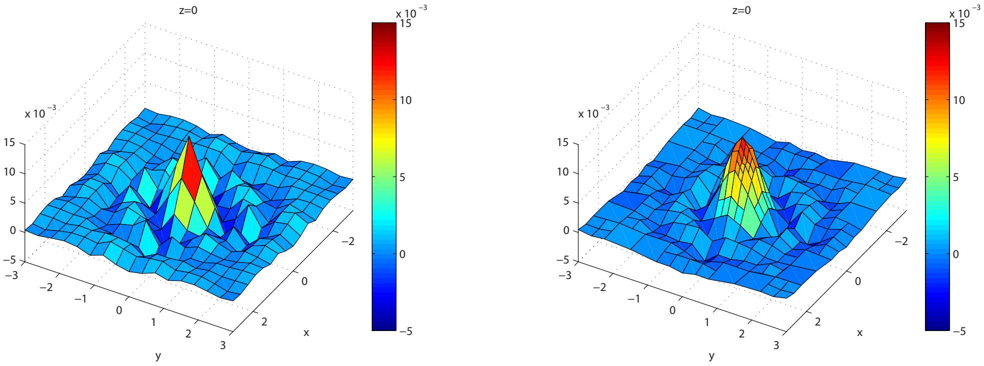
(a) The baseline image.

(b) The image reconstructed using the uniform meshes for both forward and inverse problems.



(c) The image reconstructed using our adaptive mesh and the uniform mesh for forward and inverse problems respectively.

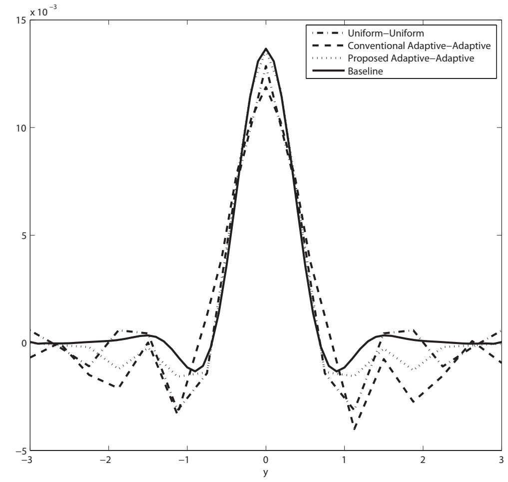
(d) The image reconstructed using our adaptive mesh for both forward and inverse problems.



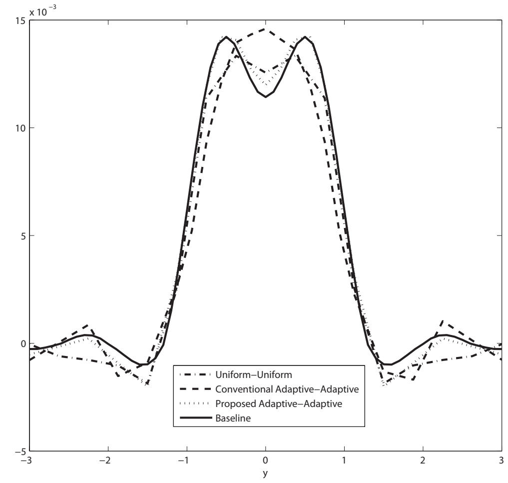
(e) The image reconstructed using the conventional adaptive mesh and the uniform mesh for forward and inverse problems respectively.

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

Figure 3.4: The reconstruction results of Simulation Study 1, Case 1, on plane $z = 0$, with $\mu_{axf} = 0.015$cm^{-1}, $r = 1$mm.



(a) The reconstructed images along the y-axis, in Simulation 1, Case 1, with $\mu_{axf} = 0.015\text{cm}^{-1}$, $r = 1\text{mm}$



(b) The reconstructed images along the y-axis, in Simulation 2, Case 3, with $\mu_{axf} = 0.015\text{cm}^{-1}$, $r = 2\text{mm}$, $d = 1\text{cm}$,

Figure 3.5: The reconstructed images μ, μ^D_{UU}, μ^D_{AA} and μ^D_{CAA} along the y-axis in Simulation Study 1, Case 1 and Simulation Study 2, Case 3.

Table 3.5: PVR of the images reconstructed by using different meshes in Simulation Study 2.

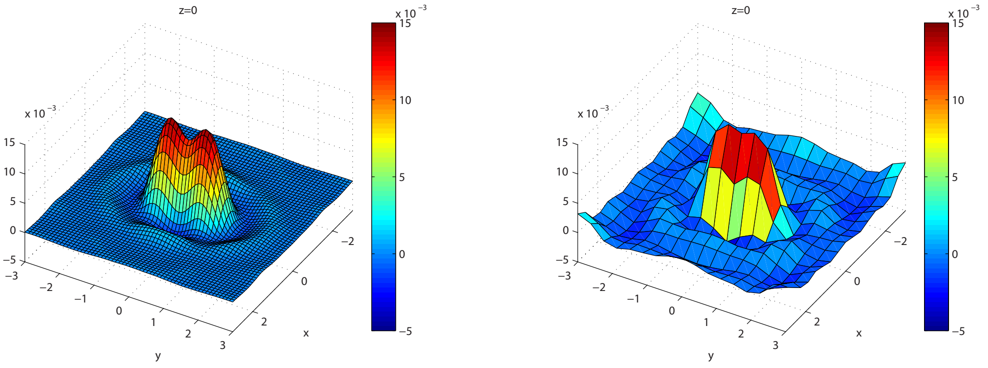
	Case 1	Case 2	Case 3	Case 4	Case 5
d (cm)	0.5	0.75	1	1.25	1.5
μ	-	-	1.227	1.957	3.638
μ^D_{UU}	-	-	1.071	1.617	3.021
μ^D_{AU}	-	-	1.136	1.818	3.529
μ^D_{AA}	-	-	1.136	1.819	3.533
μ^D_{CAU}	-	-	-	1.175	1.723
μ^D_{CAA}	-	-	-	1.184	1.730

$z = 0$ plane when the distance between the two fluorophore heterogeneities is 1 cm. Figure 3.6(a) shows the cross-section of the baseline image μ, where the two fluorophore heterogeneities are clearly distinguishable. In Figures 3.6(b), 3.6(c) and 3.6(d), we can also clearly see the two peaks of the fluorophore concentration, while in the cross-sections of μ^D_{CAU} and μ^D_{CAA} shown in Figures 3.6(e) and 3.6(f), these two peaks appear merged together.

For a close examination, in Figure 3.5(b), we also present the reconstructed images along the y-axis on $z = 0$ plane. The cross-section of μ^D_{CAA} reveals that the conventional adaptive meshing method fails to produce an image that can resolve the two distinct fluorophore heterogeneities (see dashed line in Figure 3.5(b)). On the other hand, μ^D_{AA} reconstructed by using the adaptive meshes generated by our algorithms, shows two distinguishable peaks of the fluorophore concentration with higher PVR than that of μ^D_{UU}. The results in this part of the simulation study indicates that our adaptive mesh generation algorithms can effectively improve the resolution of the reconstructed images as compared to other meshing schemes.

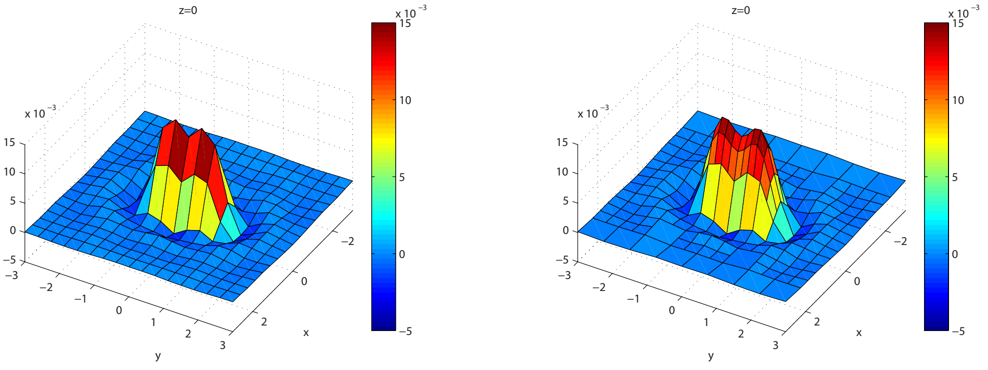
3.7.3.4 Convergence Rate

To study the convergence of the discretized inverse problem solution to the exact inverse problem solution when different types of meshes are used in FDOT reconstruction, we solved both forward and inverse problems on a series of meshes with gradually increasing discretization levels in each reconstruction scenario. Note that, for each discretization level, the mesh for the forward problem is always slightly



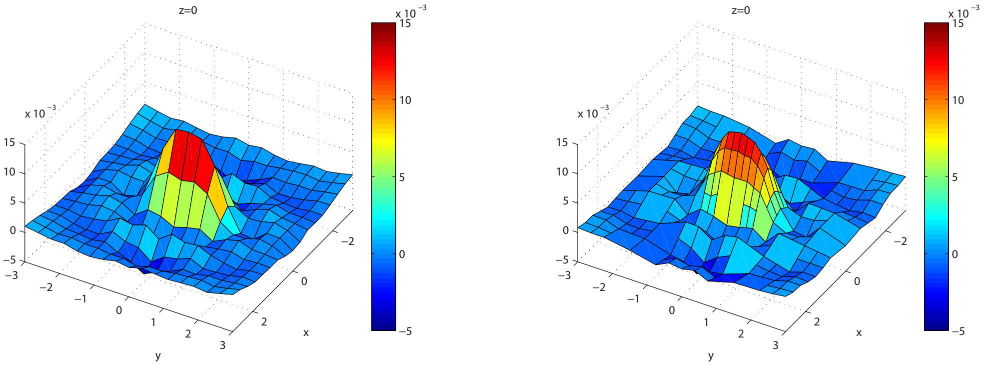
(a) The baseline image.

(b) The image reconstructed using the uniform meshes for both forward and inverse problems.



(c) The image reconstructed using our adaptive mesh and the uniform mesh for forward and inverse problems respectively.

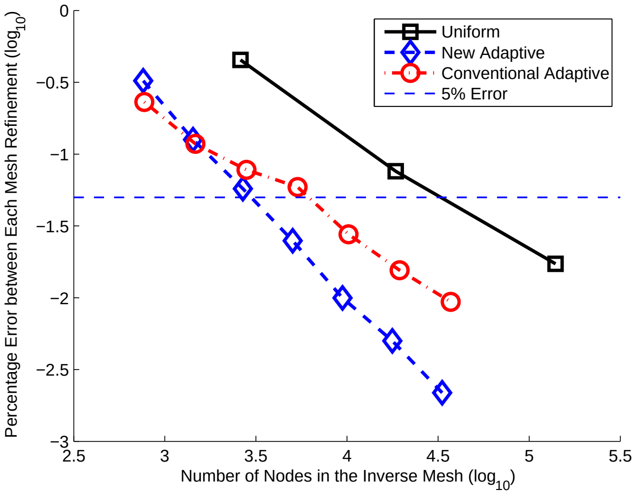
(d) The image reconstructed using our adaptive mesh for both forward and inverse problems.



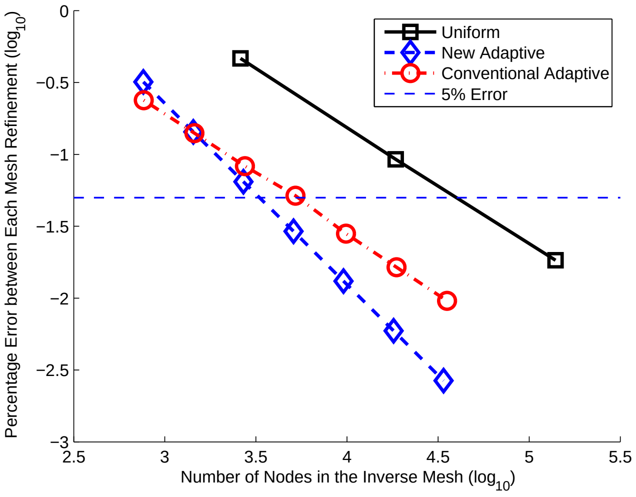
(e) The image reconstructed using the conventional adaptive mesh and the uniform mesh for forward and inverse problems respectively.

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

Figure 3.6: The reconstruction results of Simulation Study 2, Case 3, on plane $z = 0$, with $\mu_{axf} = 0.015$cm^{-1}, $r = 2$ mm, $d = 1$ cm.



(a) The change in percent $H^1(\Omega)$ norm of error versus the number of nodes used in the inverse problem mesh in Simulation 1, Case 1, with $\mu_{axf} = 0.015\text{cm}^{-1}$, $r = 1\text{mm}$



(b) The change in percent $H^1(\Omega)$ norm of error versus the number of nodes used in the inverse problem mesh in Simulation 2, Case 3, with $\mu_{axf} = 0.015\text{cm}^{-1}$, $r = 2\text{mm}$, $d = 1\text{cm}$,

Figure 3.7: The change in percent $H^1(\Omega)$ norm of error versus the number of nodes used in the inverse problem mesh in Simulation Study 1, Case 1 and Simulation Study 2, Case 3. The horizontal dash line indicates the 5% error level between each mesh refinement.

higher than the mesh for the inverse problem (there are more nodes in the mesh for the forward problem than in the mesh for the inverse problem). At the same time, we calculated the $H^1(\Omega)$ norm of the difference between the two images reconstructed on two consecutive discretization levels, and normalized it with the $H^1(\Omega)$ norm of the image reconstructed on the lower discretization level. We refer to this as the percentage error between the mesh refinement. Clearly, this error is bounded by the sum of the percentage error due to discretization in these two images, therefore it decreases as the discretization level increases. We note that this error, which reflects the gain of the image reconstruction accuracy that can be obtained by a further mesh refinement, can be used as an indicator for the error due to discretization in the reconstructed image.

Figure 3.7 shows the change in the $H^1(\Omega)$ norm of error versus the number of nodes used in the inverse problem mesh generated by different adaptive mesh generation algorithms for Simulation Study 1, Case 1 and Simulation Study 2, Case 3. The x-axis represents the number of nodes in the inverse mesh, and the y-axis represents the error. Note that both quantities are in logarithmic scale. We observe that our adaptive mesh generation algorithms have the fastest convergence rate as compared to the conventional and uniform meshing schemes. The horizontal dash lines in Figure 3.7 represent the 5% error between each mesh refinement. When the convergence curves fall below those 5% error lines, the further increased discretization level will reduce the error less than 5% of the total $H^1(\Omega)$ norm of the reconstructed image, and we consider that discretization level is high enough to obtain an relatively accurate image. Figure 3.7 shows that, to achieve this accuracy level, one needs to use the uniform mesh with over 30,000 nodes and the conventional adaptive mesh with over 6,000 nodes for the inverse problem. On the other hand, it only requires roughly 3,500 nodes in the mesh generated by our algorithm.

3.8 Conclusion

Based on the two error bounds derived in the previous chapter, we developed novel adaptive mesh generation algorithms for the forward and inverse problems of FDOT. These adaptive mesh generation algorithms aim to reduce the error due to discretization in the reconstructed optical images while constraining the number of discretization points in the discretized problems. The key feature that distinguishes our algorithms from the existing conventional adaptive meshing method is that our algorithms take into account the interdependence of FDOT forward and inverse problems. This implies that the solution of one problem will affect the discretization of the other problem in our algorithms. On the other hand, the conventional method always independently discretizes these two problems. In addition, we analytically showed that the conventional adaptive meshing method relying on the interpolation error bounds may lead to higher error in the reconstructed images than expected.

We illustrated the mesh refinement process of our adaptive mesh generation algorithms using a toy problem and also briefly discussed the computational complexities. We showed that, by using the appropriate approximations, the complexities of our algorithms can be reduced from $\mathcal{O}(N_\Delta^i N_D)$ and $\mathcal{O}(N_\Delta^j N_S)$ to $\mathcal{O}(N_\Delta^i)$ and $\mathcal{O}(N_\Delta^j)$ for the forward problem, and reduced from $\mathcal{O}(N_\Delta N_S N_D)$ to $\mathcal{O}(N_\Delta)$ for the inverse problem. In this case, the computational complexities of our algorithms are equal to that of the conventional method.

Finally, our simulation studies showed that using the adaptive meshes generated by the conventional method may lead to higher errors in the reconstructed images, especially when the fluorophore heterogeneity is small. As compared to the uniform and conventional methods, given a limited number of discretization points in the discretized forward and inverse problems, our algorithms can significantly improve the accuracy of the reconstructed image in terms of the reduced discretization error, improved small target detectability, image resolution, and convergence rate.

CHAPTER 4
EFFECT OF DISCRETIZATION IN FDOT IN THE PRESENCE OF MEASUREMENT NOISE

4.1 Introduction

In Chapter 2, we analyzed the effect of forward and inverse problem discretizations on the accuracy of FDOT reconstruction under the assumption that the measurements are noise-free. This assumption simplifies the error analysis and the development of the adaptive mesh generation algorithms for FDOT. However the important aspects of our results can also be extended to a more general measurement model with noise involved. In this chapter, we consider the measurements contaminated with additive noise and analyze the discretization error in this situation. We still focus on the estimation of the fluorophore concentration using the CW method, model the forward problem by a pair of coupled diffusion equations, and solve them by the FEMs. For the inverse problem, we first assume that the noise is an independent multivariate Gaussian noise and the fluorophore concentration is a Gaussian random field. Based on these assumptions, we formulate the inverse problem as an optimization problem in the form of MAP estimator of the fluorophore concentration at each linearized iteration. We also provide an inverse problem formulation in the form of ML estimator in the case when the *a priori* information on fluorophore concentration is not available. Following the approach we presented in Chapter 2, we transform both optimization problems into variational problems. For the inverse problem formulated as ML estimator, we regularize it with zero-th order Tikhonov regularization to address its ill-posedness. Finally, we discretize and solve these variational problems with the Galerkin projection method.

Since the measurements are contaminated with random noise, both inverse problem solution as well as error due to discretization are random fields, rather than the deterministic functions as we considered in the previous chapters. To quantitatively assess the error due to discretization, we define the MSE between the exact solution and the discretized solution of the inverse problem, and use this as the figure of merit to evaluate the accuracy of the FDOT reconstruction. Then we analyze this MSE and derive new error bounds that present a direct relationship among the MSE, the forward and inverse problem discretizations, the measurement noise, and the *a priori* information on fluorophore concentration. As compared to the error bounds given in Chapter 2, the new error bounds not only include the factors identified previously (such as the number of sources and detectors, their configuration and their positions with respect to the fluorophore heterogeneity), but also take into account the noise statistics as well as the *a priori* information on fluorophore concentration. Based on these new error bounds, we develop new adaptive mesh generation algorithms for the FDOT forward and inverse problems in the next chapter.

4.2 Models for Measurement Noise and Fluorophore Concentration

When the measurements are contaminated with additive noise, the measurement model in (2.17) can be extended in a straight forward manner:

$$\begin{aligned} \mathbf{\Gamma} &= \mathcal{A}\mu + \boldsymbol{\varepsilon} \\ &= \mathbf{\Gamma}_0 + \boldsymbol{\varepsilon}, \end{aligned} \tag{4.1}$$

where $\boldsymbol{\varepsilon} = [\varepsilon_{1,1}, \ldots, \varepsilon_{1,N_D}, \varepsilon_{2,1}, \ldots, \varepsilon_{N_S,N_D}]^T$ is the noise vector and $\mathbf{\Gamma}_0 = \mathcal{A}\mu$ is the vector of ideal noise-free measurements. Without loss of generality, we assume the components of the noise vector are mutually statistically independent Gaussian random variables with zero-mean and known variance $\sigma^2_{\boldsymbol{\varepsilon},ij} \geq 0$, for $i = 1, \ldots, N_S$ and $j = 1, \ldots, N_D$. Thus, the covariance matrix of $\boldsymbol{\varepsilon}$ is given by

$$\mathbf{\Sigma}_{\boldsymbol{\varepsilon}} = \operatorname{diag}\left([\sigma^2_{\boldsymbol{\varepsilon},11}, \ldots, \sigma^2_{\boldsymbol{\varepsilon},1N_D}, \sigma^2_{\boldsymbol{\varepsilon},21}, \ldots, \sigma^2_{\boldsymbol{\varepsilon},N_SN_D}]^T\right), \tag{4.2}$$

where diag$(\cdot)$ denotes a diagonal matrix in which the k^{th} element along the diagonal is the k^{th} component of the vector inside the bracket.

We model the fluorophore concentration image as a Gaussian random field and assume that it is statically independent of the additive noise. Particularly, we assume that the fluorophore concentration μ has mean $\mu_0 = \eta\mu_{axf}^0$, which is the known background fluorophore concentration when we formulate the Born approximation, and that $\mu(\boldsymbol{r})$ and $\mu(\acute{\boldsymbol{r}})$, $\boldsymbol{r} \neq \acute{\boldsymbol{r}}$, are mutually statistically independent. Thus, we define

$$\begin{aligned}
\mathrm{E}[\mu(\boldsymbol{r})] &= \mu_0(\boldsymbol{r}), && (4.3)\\
\mathrm{Cov}_{\mu\mu}(\boldsymbol{r},\acute{\boldsymbol{r}}) &= \mathrm{E}\left[[\mu(\boldsymbol{r})-\mu_0(\boldsymbol{r})][\mu(\acute{\boldsymbol{r}})-\mu_0(\acute{\boldsymbol{r}})]\right] \\
&=: \kappa(\boldsymbol{r})\delta(\boldsymbol{r}-\acute{\boldsymbol{r}}), && (4.4)
\end{aligned}$$

where E denotes expectation and $\kappa(\boldsymbol{r}) \geq 0$ is the variance of $\mu(\boldsymbol{r})$. Note that, in the $(k+1)^{\mathrm{th}}$ iteration of the iterative reconstruction, we assume $\hat{\mu}_{k+1}$ has mean $\hat{\mu}_k$, which is the estimate obtained at the k^{th} iteration.

4.3 Inverse Problem Formulation as MAP estimator

Based on the noise and *a priori* fluorophore concentration models, we first consider the MAP estimator for μ which is given by the following constrained minimization problem [100] in Bayesian framework:

$$\begin{aligned}
\hat{\mu} &= \min_{\mu\in L^2(\Omega)} J_{MAP}(\mu) \\
&= \min_{\mu\in L^2(\Omega)} \left[J_{LH}(\mu) + J_{PR}(\mu)\right], && (4.5)
\end{aligned}$$

where we seek a solution $\hat{\mu} \in L^2(\Omega)$ that minimize the functional J_{MAP} representing the posteriori probability of fluorophore concentration. In (4.5), the functional J_{LH} corresponds to the likelihood of measurement $\boldsymbol{\Gamma}$ given fluorophore concentration μ. J_{LH} is further given by

$$\begin{aligned}
J_{LH}(\mu) &= \left[\boldsymbol{\Gamma} - \mathcal{A}\mu\right]^T \boldsymbol{\Sigma}_{\boldsymbol{\mathcal{E}}}^{-1} \left[\boldsymbol{\Gamma} - \mathcal{A}\mu\right] \\
&= \sum_{i,j}^{N_S,N_D} \frac{1}{\sigma_{\boldsymbol{\mathcal{E}},ij}^2} \left[\Gamma_{i,j} - (\mathcal{A}\mu)_{i,j}\right]^2,
\end{aligned}$$

where

$$\boldsymbol{\Sigma}_{\boldsymbol{\mathcal{E}}}^{-1} = \mathrm{diag}\left(\left[\frac{1}{\sigma^2_{\boldsymbol{\mathcal{E}},11}}, \dots, \frac{1}{\sigma^2_{\boldsymbol{\mathcal{E}},1N_D}}, \frac{1}{\sigma^2_{\boldsymbol{\mathcal{E}},21}}, \dots, \frac{1}{\sigma^2_{\boldsymbol{\mathcal{E}},N_S N_D}}\right]^T\right),$$

is also a diagonal matrix. On the other hand, the *a priori* model of fluorophore concentration is incorporated by the functional J_{PR} given by

$$J_{PR}(\mu) \;=\; \int_\Omega \frac{1}{\kappa(\boldsymbol{r})}\left[\mu(\boldsymbol{r}) - \mu_0(\boldsymbol{r})\right]^2 d\boldsymbol{r}.$$

Taking the Gâteaux derivative of (4.5) with respect to μ, and setting it equal to zero yields:

$$\left(\mathcal{A}^* \boldsymbol{\Sigma}_{\boldsymbol{\mathcal{E}}}^{-1} \mathcal{A}\mu_{MAP}\right)(\boldsymbol{r}) + \frac{\mu_{MAP}(\boldsymbol{r})}{\kappa(\boldsymbol{r})} \;=\; \left(\mathcal{A}^* \boldsymbol{\Sigma}_{\boldsymbol{\mathcal{E}}}^{-1} \boldsymbol{\Gamma}\right)(\boldsymbol{r}) + \frac{\mu_0(\boldsymbol{r})}{\kappa(\boldsymbol{r})}, \tag{4.6}$$

where μ_{MAP} is the solution of the optimization problem (4.5). We further define $\mathcal{B}_{LH} : L^2(\Omega) \to L^2(\Omega)$ as

$$\begin{aligned} (\mathcal{B}_{LH}\mu)(\boldsymbol{r}) \;&:=\; (\mathcal{A}^* \boldsymbol{\Sigma}_{\boldsymbol{\mathcal{E}}}^{-1} \mathcal{A}\mu)(\boldsymbol{r}) \\ &= \int_\Omega \sum_{i,j}^{N_S,N_D} \frac{1}{\sigma^2_{\boldsymbol{\mathcal{E}},ij}} a^*_{ij}(\boldsymbol{r}) a_{ij}(\acute{\boldsymbol{r}})\mu(\acute{\boldsymbol{r}}) d\acute{\boldsymbol{r}}, \end{aligned} \tag{4.7}$$

and express (4.6) as follows:

$$\left(\mathcal{B}_{LH}\mu_{MAP}\right)(\boldsymbol{r}) + \frac{\mu_{MAP}(\boldsymbol{r})}{\kappa(\boldsymbol{r})} \;=\; \left(\mathcal{A}^* \boldsymbol{\Sigma}_{\boldsymbol{\mathcal{E}}}^{-1} \boldsymbol{\Gamma}\right)(\boldsymbol{r}) + \frac{\mu_0(\boldsymbol{r})}{\kappa(\boldsymbol{r})}. \tag{4.8}$$

We use the Galerkin method [86] to solve the integral equation defined in (4.8). Thus, we first define the variational form of (4.8) with a similar bilinear form as we used in Chapter 2:

$$\mathcal{F}_{MAP}(\psi, \mu_{MAP}) \;=\; (\psi, \mathcal{A}^* \boldsymbol{\Sigma}_{\boldsymbol{\mathcal{E}}}^{-1} \boldsymbol{\Gamma}) + (\psi, \frac{\mu_0}{\kappa}), \quad \forall \psi \in L^2(\Omega), \tag{4.9}$$

where

$$\mathcal{F}_{MAP}(\psi,\mu) \;:=\; (\psi,\mathcal{B}_{LH}\mu)+(\psi,\frac{\mu}{\kappa}). \tag{4.10}$$

We note that, the bilinear form $\mathcal{F}_{MAP}$ is bounded and coercive when appropriate κ is chosen (see Appendix C). Then, by Lax-Milgram lemma, a unique solution for the inverse problem (4.9) exists.

4.4 Inverse Problem Formulation as ML estimator

When the *a priori* information on the fluorophore concentration is not available, one can formulate the inverse problem as the ML estimator of fluorophore concentration with $J_{PR}(\mu)=0$ in (4.5). In this respect, the inverse problem is formulated by the following optimization problem where only the log-likelihood function of the measurement is minimized [101]:

$$\begin{aligned}\hat{\mu} &= \min_{\mu\in L^2(\Omega)} J_{LH}(\mu)\\ &= \min_{\mu\in L^2(\Omega)} \sum_{i,j}^{N_S,N_D} \frac{1}{\sigma^2_{\mathcal{E},ij}}\left[\Gamma_{i,j}-(\mathcal{A}\mu)_{i,j}\right]^2 .\end{aligned} \tag{4.11}$$

Similarly, taking the Gâteaux derivative of (4.11) with respect to μ, we obtain an integral equation that the solution of the optimization problem μ_{ML} must satisfy:

$$\left(\mathcal{B}_{LH}\mu_{ML}\right)(\boldsymbol{r}) \;=\; \left(\mathcal{A}^*\boldsymbol{\Sigma}_{\boldsymbol{\mathcal{E}}}^{-1}\boldsymbol{\Gamma}\right)(\boldsymbol{r}),$$

Since both $\mathcal{A}$ and $\mathcal{A}^*$ are compact, the inverse problem (4.12) is ill-posed. Therefore, we consider the following regularized form of (4.12):

$$\left((\mathcal{B}_{LH}+\lambda\mathcal{I})\mu_{ML}\right)(\boldsymbol{r}) \;=\; \left(\mathcal{A}^*\boldsymbol{\Sigma}_{\boldsymbol{\mathcal{E}}}^{-1}\boldsymbol{\Gamma}\right)(\boldsymbol{r}), \tag{4.12}$$

where $\mathcal{I}: L^2(\Omega)\rightarrow L^2(\Omega)$ is an identity operator, and λ is a small positive constant. Similar to Chapter 2, we assume that λ is appropriately chosen.

Then, the variational form of the ML estimator of fluorophore concentration is given by

$$\mathcal{F}_{ML}(\psi, \mu_{ML}) \;=\; (\psi, \mathcal{A}^* \boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}^{-1} \boldsymbol{\Gamma}), \quad \forall \psi \in L^2(\Omega), \tag{4.13}$$

where

$$\mathcal{F}_{ML}(\psi, \mu) \;:=\; (\psi, \mathcal{B}_{LH}\mu) + \lambda(\psi, \mu), \tag{4.14}$$

Clearly, the inverse problem formulation of the ML estimator for fluorophore concentration is a special case of the MAP estimator when

$$\kappa(\boldsymbol{r}) = \frac{1}{\lambda}, \qquad \mu_0(\boldsymbol{r}) = 0.$$

It also can be shown that the bilinear form $\mathcal{F}_{ML}$ is bounded and coercive given $\lambda > 0$(see Appendix C), and a unique solution for the inverse problem (4.13) exists.

4.5 Discretization of Inverse Problem

The forward problem discretization remains same as the one we discussed in Chapter 2. Following the discretization scheme for the forward problem in Chapter 2, we solve them and obtain the finite-dimensional solutions Φ_i and G_j^*, as well as the approximated operators $\widetilde{\mathcal{A}}$ and $\widetilde{\mathcal{A}}^*$. Then, substituting them into (4.8), we obtain an approximation $\widetilde{\mathcal{B}}_{LH}$ to $\mathcal{B}_{LH}$ defined as

$$\begin{aligned}(\widetilde{\mathcal{B}}_{LH}\mu)(\boldsymbol{r}) &= \left(\widetilde{\mathcal{A}}^* \boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}^{-1} \widetilde{\mathcal{A}}\mu\right)(\boldsymbol{r}) \\ &= \int_\Omega \sum_{i,j}^{N_S,N_D} \frac{1}{\sigma^2_{\boldsymbol{\varepsilon},ij}} A^*_{ij}(\boldsymbol{r}) A_{ij}(\acute{\boldsymbol{r}})\mu(\acute{\boldsymbol{r}}) d\acute{\boldsymbol{r}}. \end{aligned} \tag{4.15}$$

Further, let $V(\Omega) \subset L^2(\Omega)$ denote a finite-dimensional subspace of dimension N, spanned by the first-order Lagrange basis functions $\{L_k, k = 1, \ldots, N\}$ which are associated with the set of points $\{\boldsymbol{r}_p, p = 1, \ldots, N\}$ on Ω. Then, we define the finite-dimensional counterparts $\Psi \in V(\Omega)$ and $\mu_{MAP}^D \in V(\Omega)$ of ψ and μ_{MAP} as

$$\Psi(\boldsymbol{r}) := \sum_{k=1}^{N} p_k L_k(\boldsymbol{r}), \tag{4.16}$$

$$\mu_{MAP}^D(\boldsymbol{r}) := \sum_{k=1}^{N} m_k L_k(\boldsymbol{r}), \tag{4.17}$$

where p_k and m_k are unknown coefficients.

Substituting $\mathcal{B}_{LH}$ with $\widetilde{\mathcal{B}}_{LH}$ and replacing ψ and μ_{MAP} with Ψ and μ_{MAP}^D, we arrive at the fully discretized inverse problem for the MAP estimator of fluorophore concentration:

$$\widetilde{\mathcal{F}}_{MAP}(\Psi, \mu_{MAP}^D) = (\Psi, \widetilde{\mathcal{A}}^* \boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}^{-1} \boldsymbol{\Gamma}) + (\Psi, \frac{\mu_0}{\kappa}), \quad \forall \Psi \in V(\Omega), \tag{4.18}$$

where

$$\widetilde{\mathcal{F}}_{MAP}(\psi, \mu) \ := \ (\psi, \widetilde{\mathcal{B}}_{LH}\mu) + (\psi, \frac{\mu}{\kappa}). \tag{4.19}$$

Similarly, we can also obtain the fully discretized inverse problem for the ML estimator of fluorophore concentration:

$$\widetilde{\mathcal{F}}_{ML}(\Psi, \mu_{ML}^D) \ = \ (\Psi, \widetilde{\mathcal{A}}^* \boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}^{-1} \boldsymbol{\Gamma}), \quad \forall \Psi \in V(\Omega), \tag{4.20}$$

and

$$\widetilde{\mathcal{F}}_{ML}(\psi, \mu) \ := \ (\psi, \widetilde{\mathcal{B}}_{LH}\mu) + \lambda(\psi, \mu), \tag{4.21}$$

where

$$\mu_{ML}^D(\boldsymbol{r}) := \sum_{k=1}^{N} m_k L_k(\boldsymbol{r}) \in V(\Omega), \tag{4.22}$$

is the finite-dimensional counterpart of μ_{ML}. We note that both (4.18) and (4.20) can be transformed into matrix equations as the one in (2.50), from which the unknown coefficients m_k, $k = 1, \dots, N$, can be solved.

4.6 Discretization Error Analysis in the Presence of Measurement Noise

In this section, we analyze the effect of forward and inverse problem discretizations on the accuracy of FDOT reconstruction in the presence of measurement noise. Due to the random nature of the measurements as well as the fluorophore concentration, the reconstructed images obtained are random fields rather than deterministic functions which we discussed in the previous chapters. We are interested in quantifying the difference between the estimates μ_{MAP} and μ_{MAP}^D or the difference between μ_{ML} and μ_{ML}^D, which represents the error due to discretization in μ_{MAP}^D or μ_{ML}^D, respectively. Thus, we define

$$\begin{aligned} e_{MAP}(\boldsymbol{r}) &:= \mu_{MAP}(\boldsymbol{r}) - \mu_{MAP}^D(\boldsymbol{r}), \\ e_{ML}(\boldsymbol{r}) &:= \mu_{ML}(\boldsymbol{r}) - \mu_{ML}^D(\boldsymbol{r}), \end{aligned}$$

and quantify these differences in terms of the MSEs defined as follows:

$$\mathrm{MSE}[\mu_{MAP}^D] := \int_\Omega \mathrm{E}\left[|e_{MAP}(\boldsymbol{r})|^2\right] d\boldsymbol{r}, \tag{4.23}$$

$$\mathrm{MSE}[\mu_{ML}^D] := \int_\Omega \mathrm{E}\left[|e_{ML}(\boldsymbol{r})|^2\right] d\boldsymbol{r}. \tag{4.24}$$

Since the ML estimate of fluorophore concentration is only a special case of the MAP estimate, in the following sections, we first analyze the MSE for the MAP estimate, then we give the results for the ML estimate case directly. For $\mathrm{MSE}[\mu_{MAP}^D]$, we can express it as

$$\mathrm{MSE}[\mu_{MAP}^D] = \mathrm{Bias}^2[\mu_{MAP}^D] + \mathrm{Var}[\mu_{MAP}^D],$$

where

$$\begin{aligned}\mathrm{Bias}^2[\mu^D_{MAP}] &:= \int_\Omega \left|\mathrm{E}\left[e_{MAP}(\boldsymbol{r})\right]\right|^2 d\boldsymbol{r}, &(4.25)\\ \mathrm{Var}[\mu^D_{MAP}] &:= \int_\Omega \mathrm{E}\left[\left|e_{MAP}(\boldsymbol{r})-\mathrm{E}[e_{MAP}(\boldsymbol{r})]\right|^2\right] d\boldsymbol{r}. &(4.26)\end{aligned}$$

We refer to $\mathrm{Bias}[\mu^D_{MAP}]$ as the bias of μ^D_{MAP} with respect to the exact estimate μ_{MAP} and $\mathrm{Var}[\mu^D_{MAP}]$ as the variance of μ^D_{MAP}. We note that $\mathrm{MSE}[\mu^D_{ML}]$ can also be decomposed into $\mathrm{Bias}^2[\mu^D_{ML}]$ and $\mathrm{Var}[\mu^D_{ML}]$ in the same manner.

In the following sections, we analyze the effect of discrtization and present two upper bounds for $\mathrm{Bias}^2[\mu^D_{MAP}]$ and $\mathrm{Var}[\mu^D_{MAP}]$, respectively. Then we discuss the implications of these error bounds for the forward and inverse problem discretizations, and give the results for $\mathrm{Bias}^2[\mu^D_{ML}]$ and $\mathrm{Var}[\mu^D_{ML}]$ in the ML estimate case. Finally, we discuss the validity of the models we use for measurement noise and fluorophore concentration, and present the results for the case involving correlated noise and *a priori* fluorophore concentration models.

4.6.1 Bias of the MAP Estimate

In the following theorem, we first present an upper bound for $\mathrm{Bias}^2[\mu^D_{MAP}]$ defined in (4.25).

Theorem 3:

Consider the Galerkin projection of the variational problems (2.34), (2.35), (4.9) described in Sections 2.5.1 and 4.5, respectively. Let $\bar{\mu}_{MAP}(\boldsymbol{r}) := \mathrm{E}[\mu_{MAP}(\boldsymbol{r})]$, then $\bar{\mu}_{MAP}$ satisfies the following variational problem:

$$\mathcal{F}_{MAP}(\psi,\bar{\mu}_{MAP}) = (\psi,\mathcal{A}^*\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}^{-1}\bar{\boldsymbol{\Gamma}}) + (\psi,\frac{\mu_0}{\kappa}), \quad \forall\psi\in L^2(\Omega),$$

where $\bar{\boldsymbol{\Gamma}} = \mathrm{E}[\boldsymbol{\Gamma}] = \mathcal{A}\mu_0$. Furthermore, $\mathrm{Bias}^2[\mu^D_{MAP}]$ satisfies the following inequality:

$$\mathrm{Bias}^2[\mu^D_{MAP}] \le C\left[B_1+B_2+B_3\right]^2, \tag{4.27}$$

where

$$\begin{aligned}
B_1 &= \sum_{i=1}^{N_S} \sum_{n,j}^{N_\Delta^i, N_D} \left(F_{ij}^1 \|g_j^* \bar{\mu}_{MAP}\|_{0,ni} + F_{ij}^2 \|g_j^*\|_{\infty,ni} \right) \|\phi_i\|_{1,ni} h_{ni}, \\
B_2 &= \sum_{j=1}^{N_D} \sum_{m,i}^{N_\Delta^{*j}, N_S} \left(F_{ij}^1 \|\phi_i \bar{\mu}_{MAP}\|_{0,mj} + F_{ij}^2 \|\phi_i\|_{\infty,mj} \right) \|g_j^*\|_{1,mj} h_{mj}, \\
B_3 &= \sum_{t=1}^{N_\Delta} \left(\sum_{i,j}^{N_S,N_D} I_{ij}^1 \|G_j^* \Phi_i\|_{0,t} + I_t^2 \right) \|\bar{\mu}_{MAP}\|_{1,t} h_t,
\end{aligned}$$

with

$$\begin{aligned}
F_{ij}^1 &= \frac{2\|\kappa\|_\infty \|g_j^* \phi_i\|_0}{\sigma_{\mathcal{E},ij}^2}, \quad F_{ij}^2 = \frac{\|\kappa\|_\infty |\bar{\Gamma}_{i,j}|}{\sigma_{\mathcal{E},ij}^2}, \\
I_{ij}^1 &= \frac{\|\kappa\|_\infty \|G_j^* \Phi_i\|_0}{\sigma_{\mathcal{E},ij}^2}, \quad I_t^2 = \|\kappa\|_\infty \left\| \frac{1}{\kappa} \right\|_{\infty,t},
\end{aligned}$$

and C is a positive constant independent of the discretization parameters h_{ni}, h_{mj} and h_t.

Proof: See Appendix H. □

Equation (4.27) in Theorem 3 presents an upper bound for $\text{Bias}^2[\mu_{MAP}^D]$, which takes into account the noise statistics and the *a priori* information of fluorophore concentration, in addition to the interdependence between the forward and inverse problem solutions. In this error bound, B_1 and B_2 represent the contribution from the forward problem discretization. To keep these quantities small, the mesh parameters h_{ni} and h_{mj} of the n^{th} and m^{th} elements in the meshes used in solving Φ_i and G_j^*, respectively, have to be chosen small when their corresponding scaling factors

$$\sum_{j=1}^{N_D} \left(F_{ij}^1 \|g_j^* \bar{\mu}_{MAP}\|_{0,ni} + F_{ij}^2 \|g_j^*\|_{\infty,ni} \right) \|\phi_i\|_{1,ni},$$

and

$$\sum_{i=1}^{N_S} \left(F_{ij}^1 \|\phi_i \bar{\mu}_{MAP}\|_{0,mj} + F_{ij}^2 \|\phi_i\|_{\infty,mj} \right) \|g_j^*\|_{1,mj},$$

are large on those elements. Further examination of these factors suggests an adaptive refinement scheme within each mesh, because $\|g_j^* \bar{\mu}_{MAP}\|_{0,ni}$, $\|g_j^*\|_{\infty,ni}$, $\|\phi_i\|_{1,ni}$, $\|\phi_i \bar{\mu}_{MAP}\|_{0,mj}$, $\|\phi_i\|_{\infty,mj}$, and $\|g_j^*\|_{1,mj}$ all vary within the mesh. This mesh refinement scheme is similar to the one suggested by Theorem 1 in Chapter 2: For the i^{th} source or the j^{th} detector, it refines the mesh close to that source or detector, as well as around the fluorophore heterogeneity and other detectors or sources. At the same time, the coefficient $F_{ij}^1 = 2\|g_j^* \phi_i\|_0 \|\kappa\|_\infty / \sigma_{\boldsymbol{\mathcal{E}},ij}^2$ and $F_{ij}^2 = |\bar{\Gamma}_{i,j}| \|\kappa\|_\infty / \sigma_{\boldsymbol{\mathcal{E}},ij}^2$ in B_1 and B_2 may vary for different source-detector pairs. To keep B_1 and B_2 low, one has to generate finer meshes for the source-detector pairs with smaller noise variances and higher magnitude of measurements, as compared to those pairs with higher noise variances and smaller measurements. In this respect, the error bound in Theorem 3 suggest a new adaptive mesh refinement scheme across different meshes in solving Φ_i and G_j^* based on the measurements and the noise statistics. We note that this is a major difference between the implications of the error bounds in this chapter and those from Chapter 2.

In B_3, which corresponds to the contribution from the inverse problem discretization, the discretization parameter h_t of the inverse mesh is not only scaled by the inverse problem solution $\|\bar{\mu}_{MAP}\|_{1,t}$, but also scaled by the finite element solutions of the forward problem, the noise variance and *a priori* information of the fluorophore concentration:

$$\sum_{i,j}^{N_S,N_D} I_{ij}^1 \|G_j^* \Phi_i\|_{0,t} + I_t^2,$$

where $I_{ij}^1 = \|\kappa\|_\infty \|G_j^* \Phi_i\|_0 / \sigma_{\boldsymbol{\mathcal{E}},ij}^2$, and $I_t^2 = \|\kappa\|_\infty \, \|1/\kappa\|_{\infty,t}$. This result also suggests a new adaptive mesh refinement criteria for the inverse problem of FDOT, not only based on the forward and inverse problem solutions, but also based on the noise statistics and *a priori* information of fluorophore concentration. More specifically, to keep B_3 low, one has to refine the mesh around the heterogeneity of fluorophore concentration, and the sources and detectors with low noise variances, as well as the region where the fluorophore concentration has low variance.

We further note that, when the measurements are noise-free, the inverse problem formulation in this chapter can be reduced to the one given in Chapter 2 with only zero-th order Tikhonov regularization ($\lambda_2 = 0$). In this respect, we set $\sigma^2_{\boldsymbol{\mathcal{E}},ij} = 1$ for $i = 1, \ldots, N_S$ and $j = 1, \ldots, N_D$, $\mu_0(\boldsymbol{r}) = \beta_1(\boldsymbol{r})$, and $\kappa(\boldsymbol{r}) = \lambda_1$. Then the upper bound for $\text{Bias}^2[\mu^D_{MAP}]$ can be reduced to the combination of two error bounds in (2.53) and (2.54). Finally, we note that, for the ML estimate μ^D_{ML}, the upper bound for $\text{Bias}^2[\mu^D_{ML}]$ can be given as in (4.27) with

$$F^1_{ij} = \frac{2\|g^*_j\phi_i\|_0}{\lambda\sigma^2_{\boldsymbol{\mathcal{E}},ij}}, \quad F^2_{ij} = \frac{|\bar{\Gamma}_{i,j}|}{\lambda\sigma^2_{\boldsymbol{\mathcal{E}},ij}}, \quad I^1_{ij} = \frac{\|G^*_j\Phi_i\|_0}{\lambda\sigma^2_{\boldsymbol{\mathcal{E}},ij}}, \quad I^2_t = 1.$$

4.6.2 Variance of the MAP Estimate

For $\text{Var}[\mu^D_{MAP}]$ define in (4.26), we present an upper bound for it in the following theorem.

Theorem 4:

Consider the Galerkin projection of the variational problems (2.34), (2.35) and (4.9) described in Sections 2.5.1 and 4.5, respectively. Let $\pi_{ij} \in L^2(\Omega)$ be the solution of the following variational problem:

$$\mathcal{F}_{MAP}(\psi, \pi_{ij}) \;=\; (\psi, \mathcal{A}^*\boldsymbol{\Sigma}^{-1}_{\boldsymbol{\mathcal{E}}}\mathbf{e}_{ij}), \quad \forall\psi \in L^2(\Omega),$$

and $\mathbf{e}_{ij} = [0,\, \ldots,\, 1,\, \ldots,\, 0]^T \in \mathbb{R}^{N_S N_D}$, where the only non-zero entry is at the $[N_D(i-1)+j]^{\text{th}}$ position. Then $\text{Var}[\mu^D_{MAP}]$ satisfies the following inequality:

$$\text{Var}[\mu^D_{MAP}] \le C\,[V_1 + V_2 + V_3]^2\,, \tag{4.28}$$

where

$$V_1 \;=\; \sum_{i=1}^{N_S}\sum_{n,j}^{N^i_\Delta, N_D}\left(F^1_{ij}\sum_{i',j'}^{N_S,N_D}\|g^*_j D_{i'j'}\pi_{i'j'}\|_{0,ni} + F^3_{ij}D_{ij}\|g^*_j\|_{\infty,ni}\right)\|\phi_i\|_{1,ni}h_{ni},$$

$$V_2 \;=\; \sum_{j=1}^{N_D}\sum_{m,i}^{N^{*j}_\Delta, N_S}\left(F^1_{ij}\sum_{i',j'}^{N_S,N_D}\|\phi_i D_{i'j'}\pi_{i'j'}\|_{0,mj} + F^3_{ij}D_{ij}\|\phi_i\|_{\infty,mj}\right)\|g^*_j\|_{1,mj}h_{mj},$$

$$V_3 = \sum_{t=1}^{N_\Delta} \left(\sum_{i,j}^{N_S,N_D} I_{ij}^1 \|G_j^* \Phi_i\|_{0,t} + I_t^2 \right) \left(\sum_{i,j}^{N_S,N_D} \|D_{ij}\pi_{ij}\|_{1,t} \right) h_t,$$

with

$$F_{ij}^1 = \frac{2\|\kappa\|_\infty \|g_j^* \phi_i\|_0}{\sigma_{\boldsymbol{\mathcal{E}},ij}^2}, \quad F_{ij}^3 = \frac{\|\kappa\|_\infty}{\sigma_{\boldsymbol{\mathcal{E}},ij}^2}, \quad I_{ij}^1 = \frac{\|\kappa\|_\infty \|G_j^* \Phi_i\|_0}{\sigma_{\boldsymbol{\mathcal{E}},ij}^2},$$

$$I_t^2 = \|\kappa\|_\infty \left\| \frac{1}{\kappa} \right\|_{\infty,t}, \quad D_{ij} = \left[\sigma_{\boldsymbol{\mathcal{E}},ij}^2 + \int_\Omega \kappa(\boldsymbol{r}) a_{ij}^*(\boldsymbol{r}) a_{ij}(\boldsymbol{r}) d\boldsymbol{r} \right]^{1/2},$$

and C is a positive constant independent of the discretization parameters h_{ni}, h_{mj} and h_t.

Proof: See Appendix I. □

Equation (4.28) in Theorem 4 shows the effect of the forward and inverse problem discretizations, and the *a priori* information of fluorophore concentration, as well as the noise on $\mathrm{Var}[\mu_{MAP}^D]$.

In this error bound, V_1 and V_2 correspond to the contribution from the forward problem discretization, and V_3 corresponds to the contribution from the inverse problem discretization. This error bound has a similar form as the one in (4.27), but $\bar{\Gamma}_{ij}$ is replaced with the standard deviation, D_{ij}, of the $(i,j)^{\text{th}}$ measurement, and $\bar{\mu}_{MAP}$ is replaced with $D_{ij}\pi_{ij}$, where π_{ij} is the image reconstructed by the imaging system using the basis vector $\mathbf{e}_{ij}$ in the measurement space $\mathbb{R}^{N_S N_D}$. This result indicates that $\mathrm{Var}[\mu_{MAP}^D]$ is independent of the fluorophore concentration, but depends explicitly on the noise statistics as well as the factors related to the imaging geometry and the background optical properties, which are incorporated into the error bound through the functions π_{ij}. More specifically, $D_{ij}\pi_{ij}$ indicates where μ_{MAP} may have high variance due to the $(i,j)^{\text{th}}$ measurement. Therefore, to keep the error bound in (4.28) low, one has to refine the mesh in the region where $D_{ij}\pi_{ij}$ has high value, in addition to the region close to the sources and detectors.

When the measurements are noise-free, it can be shown that the bound in (4.28) equals to zero, and therefore $\mathrm{Var}[\mu_{ML}^D]$ vanishes (see Appendix I). This result is consistent with the fact that the error due to discretization is deterministic in the noise-free situation. Finally, we give the upper bound for $\mathrm{Var}[\mu^D_{ML}]$ of the ML estimate μ^D_{ML}, which is in the same form as in (4.28) but with

$$F^1_{ij} = \frac{2\|g^*_j\phi_i\|_0}{\lambda\sigma^2_{\boldsymbol{\varepsilon},ij}}, \quad F^3_{ij} = \frac{1}{\lambda\sigma^2_{\boldsymbol{\varepsilon},ij}}, \quad I^1_{ij} = \frac{\|G^*_j\Phi_i\|_0}{\lambda\sigma^2_{\boldsymbol{\varepsilon},ij}}, \quad I^2_t = 1, \quad D_{ij} = \sigma_{\boldsymbol{\varepsilon},ij}.$$

4.6.3 Total MSE of the MAP Estimate

Both error bounds in Theorems 3 and 4 present a tradeoff between minimizing $\mathrm{Bias}^2[\mu^D_{MAP}]$ and $\mathrm{Var}[\mu^D_{MAP}]$, when discretizing the forward and inverse problems. One can use either error bound to adaptively refine the mesh that can minimize $\mathrm{Bias}^2[\mu^D_{MAP}]$ or $\mathrm{Var}[\mu^D_{MAP}]$. However, minimizing $\mathrm{Bias}^2[\mu^D_{MAP}]$ or $\mathrm{Var}[\mu^D_{MAP}]$ may lead to an increase of the other. To address this tradeoff, we can combine the error bounds in Theorems 3 and 4 to obtain a total upper bound for $\mathrm{MSE}[\mu^D_{MAP}]$:

$$\mathrm{MSE}[\mu^D_{MAP}] \le C\left[M_1 + M_2 + M_3\right]^2, \tag{4.29}$$

and

$$\begin{aligned}
M_1 &= \sum_{i=1}^{N_S}\sum_{n,j}^{N^i_\Delta, N_D}\left(F^1_{ij}\left(\|g^*_j\bar{\mu}_{MAP}\|_{0,ni} + \sum_{i',j'}^{N_S,N_D}\|g^*_j D_{i'j'}\pi_{i'j'}\|_{0,ni}\right)\right.\\
&\quad \left. + \left(F^2_{ij} + F^3_{ij}D_{ij}\right)\|g^*_j\|_{\infty,ni}\right)\|\phi_i\|_{1,ni}h_{ni},\\
M_2 &= \sum_{j=1}^{N_D}\sum_{m,i}^{N^{*j}_\Delta, N_S}\left(F^1_{ij}\left(\|\phi_i\bar{\mu}_{MAP}\|_{0,mj} + \sum_{i',j'}^{N_S,N_D}\|\phi_i D_{i'j'}\pi_{i'j'}\|_{0,mj}\right)\right.\\
&\quad \left. + \left(F^2_{ij} + F^3_{ij}D_{ij}\right)\|\phi_i\|_{\infty,mj}\right)\|g^*_j\|_{1,mj}h_{mj},\\
M_3 &= \sum_{t=1}^{N_\Delta}\left(\sum_{i,j}^{N_S,N_D} I^1_{ij}\|G^*_j\Phi_i\|_{0,t} + I^2_t\right)\left(\|\bar{\mu}_{MAP}\|_{1,t} + \sum_{i,j}^{N_S,N_D}\|D_{ij}\pi_{ij}\|_{1,t}\right)h_t,
\end{aligned}$$

where F^1_{ij}, F^2_{ij}, F^3_{ij}, I^1_{ij}, I^2_t and D_{ij} are defined in Theorems 3 and 4. Note that (4.29) also holds for $\mathrm{MSE}[\mu^D_{ML}]$ with the coefficients defined for ML estimate in the previous sections.

4.7 Discussion on Model Validity

In this section, we briefly discuss the validity of our assumptions on the measurement noise and *a priori* fluorophore concentration models. In optical imaging, the measurement is the light intensity measured by the optical detector at the tissue boundary. The output signal of the optical detector is proportional to the number of photons detected per unit time, therefore measurement noise is dominated by the shot-noise, when the thermal noise is well controlled by appropriately cooling the detector. When a sufficiently large number of photons are detected, the Poisson distribution of the signal can be approximated by a Gaussian distribution with the mean equal to the average signal intensity and the variance proportional to its mean [102]. In this case, the measurement noise can be modeled by the Gaussian random variable as we assumed in Section 4.2, and the variance, $\sigma^2_{\boldsymbol{\mathcal{E}},,ij}$, $i = 1, \ldots, N_S$ and $j = 1, \ldots, N_D$, of each noise component is given by

$$\sigma^2_{\boldsymbol{\mathcal{E}},ij} = \alpha |\Gamma_{0,ij}|, \tag{4.30}$$

where $\Gamma_{0,ij}$ is the noise-free measurement obtained by the j^{th} detector due to the i^{th} source. This Gaussian approximation is widely used in CW [21, 22] and FD [26, 27, 100] imaging systems to model the measurement noise, due to sufficient number of photons can be detected in such systems. However, in the TD measurement system, due to the high sampling frequency in the time domain, the signal measured in each time interval is much lower than those measured by CW and FD systems. Therefore, Poisson distribution is a more appropriate model for measurement noise for the TD systems [28, 30].

To regularize the ill-posed FDOT inverse problem, we assume an independent Gaussian *a priori* model for the fluorophore concentration in the image reconstruction. This Gaussian model or the equivalent L^2 norm regularization in Chapter 2 are widely used in the imaging inverse problems to address their ill-posedness and incorporate *a priori* information, such as in DOT [103–105], FDOT [27, 106–108], and EIT [56]; and they are also used to model images in the image restoration problem [90, 109–111]. The Gaussian and L^2 norm regularization terms in the objective functional given by (4.5) and (2.19) lead to relatively simple optimization problems in computational respect [112]. However, it has been reported that the Gaussian model or the L^2 norm regularization may cause edge-smoothing effect in the image reconstruction and restoration. To overcome this drawback, other models or regularization functionals were proposed to model images in [112–116]. These models were shown to be effective to preserve the edges in the image, while suppressing the noise. However, these models also increase the computational requirements of the corresponding image reconstruction and restoration problems.

We note that the error analysis as well as the adaptive mesh generation algorithms presented in this thesis can also be extended to the case when other measurement noise and *a priori* fluorophore concentration models are used in FDOT reconstruction. In this case, the FDOT inverse problem formulated as the optimization problem will have a different objective functional. More specifically, J_{LH} and J_{PR} in (4.5) needs to be formulated according to the particular noise and *a priori* fluorophore concentration models used, respectively. After transforming the optimization problem into its variational form which can be solved by the FEMs, we can analyze the MSE due to discretization by decomposing it similarly into the bias and variance of the reconstructed image. The bias will depend on the mean of the measurements, as well as the mean of the fluorophore concentration estimate. On the other hand, the variance arises from the measurement noise, and can be related to the second-order statistics of the measurement noise through the images π_{ij}. By taking both forward and inverse problem discretizations into account, both upper bounds for the bias and variance can be derived as functions of the forward and inverse problem solutions, discretization parameters, and noise statistics, as well as the *a priori* information on fluorophore concentration.

Finally, we note that, although we have been focusing on the error analysis in the independent Gaussian model case, our results can be directly extended to the case involving more general Gaussian models for both measurement noise and fluorophore concentration. For a more general second-order statistic of the *a priori* Gaussian model for the fluorophore concentration, we consider $\mathrm{Cov}_{\mu\mu}(\boldsymbol{r}, \acute{\boldsymbol{r}}) = \kappa(\boldsymbol{r}, \acute{\boldsymbol{r}})$ as the kernel of a positive definite operator $\mathcal{K} : L^2(\Omega) \to L^2(\Omega)$. Similarly, for a general Gaussian noise model, we assume that the covariance matrix, $\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}$, of the measurement noise is a positive definite matrix, that is not necessarily diagonal. Then, the variational form of the inverse problem formulation becomes

$$\mathcal{F}_{MAP}(\psi, \mu_{MAP}) \;=\; \left(\psi, \mathcal{A}^*\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}^{-1}\boldsymbol{\Gamma}\right) + (\psi, \mathcal{K}^{-1}\mu_0),$$

where

$$\mathcal{F}_{MAP}(\psi, \mu) \;:=\; (\psi, \mathcal{B}_{LH}\mu) + (\psi, \mathcal{K}^{-1}\mu).$$

In this case, it can be shown that the upper bounds for $\mathrm{Bias}^2[\mu_{MAP}^D]$ and $\mathrm{Var}[\mu_{MAP}^D]$ are also of the same forms as in (4.27) and (4.28), but with new coefficients given by

$$\begin{aligned}
F_{ij}^1 &= \frac{2}{\|\kappa^{-1}\|_\infty} \sum_{p=1}^{N_S N_D} \left(\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}^{-1}\right)_{p,(i-1)N_D+j} \|g_j^*\phi_i\|_0, \\
F_{ij}^2 &= \frac{1}{\|\kappa^{-1}\|_\infty} \sum_{p=1}^{N_S N_D} \left(\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}^{-1}\right)_{p,(i-1)N_D+j} |\bar{\Gamma}_{i,j}|, \\
F_{ij}^3 &= \frac{1}{\|\kappa^{-1}\|_\infty} \sum_{p=1}^{N_S N_D} \left(\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}^{-1}\right)_{p,(i-1)N_D+j}, \\
I_{ij}^1 &= \frac{1}{\|\kappa^{-1}\|_\infty} \sum_{p=1}^{N_S N_D} \left(\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}^{-1}\right)_{p,(i-1)N_D+j} \|G_j^*\Phi_i\|_0, \\
I_t^2 &= \frac{1}{\|\kappa^{-1}\|_\infty} \|\kappa_0^*\|_0 \, \|\kappa_0\|_{\infty,t}, \\
D_{ij} &= \left[(\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}})_{(i-1)N_D+j,(i-1)N_D+j} + \int_\Omega \int_\Omega \kappa(\boldsymbol{r}, \acute{\boldsymbol{r}}) a_{ij}^*(\boldsymbol{r}) a_{ij}(\acute{\boldsymbol{r}}) d\boldsymbol{r} d\acute{\boldsymbol{r}} \right]^{1/2},
\end{aligned}$$

where κ^{-1} denote the kernel of the operator $\mathcal{K}^{-1} : L^2(\Omega) \to L^2(\Omega)$, i.e.,

$$(\mathcal{K}^{-1}\mu)(\boldsymbol{r}) = \int_\Omega \kappa^{-1}(\boldsymbol{r}, \acute{\boldsymbol{r}})\mu(\acute{\boldsymbol{r}}) d\acute{\boldsymbol{r}}, \tag{4.31}$$

and $\kappa_0(\boldsymbol{r})$ is the Kolmogorov decomposition of $\kappa^{-1}(\boldsymbol{r}, \acute{\boldsymbol{r}})$ [117]; $(\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}})_{p,q}$ and $\left(\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}^{-1}\right)_{p,q}$, $p, q = 1, \ldots, N_S N_D$, denote the entries on the p^{th} row and the q^{th} column of $\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}$ and

$\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}^{-1}$, respectively. Clearly, these coefficients reduce to those in Theorems 3 and 4 when the independent noise and fluorophore concentration models are considered in FDOT reconstruction.

4.8 Conclusion

In this chapter, we extended our analysis for the noise-free situation in Chapter 2 into the case when the measurements are corrupted by additive noise. When the additive measurement noise is involved in the FDOT reconstruction, the reconstructed image as well as the error is a random field rather than a deterministic function. In this respect, one has to seek for a different figure of merit to quantitatively assess the error due to discretization and the reconstruction accuracy than the $H^1(\Omega)$ norm of the discretization error used in the previous chapters. In this chapter, we defined the MSE due to discretization in the discretized inverse problem solution to evaluate the reconstruction accuracy in the mean square sense. In our analysis, we identified two components in the total MSE, namely the bias and the variance of the fluorophore concentration estimate. Then we analyzed these two parts and presented two new error bounds for them respectively. These two error bounds show that the MSE due to discretization is not only affected by the factors previously identified in Theorems 1 and 2, but also by the statistics of the measurement noise and the *a priori* information on fluorophore concentration. Based on these two new error bounds, in the next chapter, we develop two new adaptive mesh generation algorithms for FDOT forward and inverse problems in the presence of measurement noise.

CHAPTER 5
ADAPTIVE MESH GENERATION FOR FDOT IN THE PRESENCE OF MEASUREMENT NOISE

5.1 Introduction

In Chapter 4, taking into account the measurement noise, we analyzed the effect of discretization on the accuracy of FDOT reconstruction and derive two new error bounds for the MSE due to discretizationin in the fluorophore concentration estimate. In this chapter, based on these two error bounds, we develop two new adaptive mesh generation algorithms for FDOT forward and inverse problems in the presence of noise. Since the ML estimator of the fluorophore concentration can be viewed as a special case of the MAP estimator, we focus on developing adaptive mesh generation algorithms for the MAP case in this chapter. In this respect, the new algorithms use MSE$[\mu_{MAP}^D]$ as the figure of merit, and additionally take the noise statistics and *a priori* information on fluorophore concentration into account to improve the reconstruction accuracy while keeping the size of the discretized problems under the allowable limits.

In the following sections, we still consider two adaptive h-refinement approaches to design our new adaptive mesh generation algorithms: The first approach adaptively refines the mesh to achieve a *predetermined* error tolerance, and the second approach limits the total number of discretization points in the discretized problems. In the design of adaptive mesh generation algorithm for the forward problem, as suggested by Theorems 3 and 4 in Chapter 4, we propose a new adaptive refinement scheme across the meshes for different sources and detectors. By doing so, the adaptive meshes generated by our new algorithm not only have spatially varying resolution within each mesh, but also have varying resolution across all the meshes. After describing our new algorithms in detail, we also briefly analyze the computational complexities of the new algorithms and address several practical issues in their implementations. At the end of this chapter, we present a series of three-dimensional numerical simulations to demonstrate the implications of the error bounds given in Chapter 4 and the practical advantages of our new algorithms for FDOT in the presence of measurement noise.

5.2 Adaptive Mesh Generation for Forward Problem

We use MSE$[\mu^D_{MAP}]$ as the figure of merit, and design our adaptive mesh generation algorithms based on its upper bound given in (4.29). For the forward problem, we aim to minimize the sum of M_1 and M_2 in (4.29), which aries from the forward problem discretization. Clearly, M_1 corresponds to the discretization of the boundary value problem (2.5)-(2.7) for each source, and M_2 corresponds to the discretization of the boundary value problem (2.9)-(2.10) for each detector.

For the first type of adaptive refinement approach, let $\tilde{\varepsilon}_f$ be the *predetermined* allowable upper bound on the sum of $M1$ and $M2$, which represents the allowable error due to the forward problem discretization. Then we define

$$\varepsilon_f := \frac{1}{\sqrt{C}} \left(\sum_{i=1}^{N_S} N^i_\Delta + \sum_{j=1}^{N_D} N^{*j}_\Delta \right)^{-1} \tilde{\varepsilon}_f, \tag{5.1}$$

where C is the positive constant in (4.29) and N^i_Δ or N^{*j}_Δ is the number of elements in the mesh used in solving Φ_i or G^*_j, respectively. Let B_{ni} and B_{mj}, $i = 1, \cdots, N_S$ and $j = 1, \cdots, N_D$, be defined as

$$\begin{aligned} B_{ni} &:= \frac{\varepsilon_f}{\sum_{j=1}^{N_D} \left(F^1_{ij} \left(\|g^*_j \bar{\mu}_{MAP}\|_{0,ni} + \sum_{i',j'}^{N_S,N_D} \|g^*_j D_{i'j'} \pi_{i'j'}\|_{0,ni} \right) + \left(F^2_{ij} + F^3_{ij} D_{ij} \right) \|g^*_j\|_{\infty,ni} \right) \|\phi_i\|_{1,ni}}, && (5.2) \\ B^*_{mj} &:= \frac{\varepsilon_f}{\sum_{i=1}^{N_S} \left(F^1_{ij} \left(\|\phi_i \bar{\mu}_{MAP}\|_{0,mj} + \sum_{i',j'}^{N_S,N_D} \|\phi_i D_{i'j'} \pi_{i'j'}\|_{0,mj} \right) + \left(F^2_{ij} + F^3_{ij} D_{ij} \right) \|\phi_i\|_{\infty,mj} \right) \|g^*_j\|_{1,mj}}. && (5.3) \end{aligned}$$

In this case, if $h_{ni} > 0$ and $h_{mj} > 0$ are chosen as

$$\begin{aligned} h_{ni} &\leq B_{ni}, \\ h_{mj} &\leq B^*_{mj}, \end{aligned}$$

then by (4.29), this implies that

$$M_1 + M_2 \leq \tilde{\varepsilon}_f.$$

The first approach of adaptive mesh refinement starts with a set of coarse uniform meshes for all the sources and detectors, and it checks the size parameter h_{ni} (or h_{mj}) for each element against B_{ni} (or B^*_{mj}). Then every element with h_{ni} (or h_{mj}) larger than B_{ni} (or B^*_{mj}) is refined, until h_{ni} (or h_{mj}) of each element is less than the corresponding bound B_{ni} (or B^*_{mj}). We note that, since B_{ni} and B^*_{mj} are functions of both measurement Γ_{ij} and noise variance $\sigma^2_{\boldsymbol{\varepsilon},ij}$, and they may vary not only within each mesh, but also across different source-detector pairs. Therefore, the resulting mesh may also have varying resolution for different source-detector pairs.

For the second adaptive refinement approach, we give two error indicators for the n^{th} and m^{th} element in the forward problem meshes associated with i^{th} source and j^{th} detector, respectively:

$$\begin{aligned}
\varepsilon^i_{MAP,f}(n) \quad := \quad & \sum_{j=1}^{N_D} \left(F^1_{ij} \left(\|g^*_j \bar{\mu}_{MAP}\|_{0,ni} + \sum_{i',j'}^{N_S,N_D} \|g^*_j D_{i'j'} \pi_{i'j'}\|_{0,ni} \right) \right. \\
& \left. + \left(F^2_{ij} + F^3_{ij} D_{ij} \right) \|g^*_j\|_{\infty,ni} \right) \|\phi_i\|_{1,ni} h_{ni} \qquad (5.4) \\
\varepsilon^j_{MAP,f}(m) \quad := \quad & \sum_{i=1}^{N_S} \left(F^1_{ij} \left(\|\phi_i \bar{\mu}_{MAP}\|_{0,mj} + \sum_{i',j'}^{N_S,N_D} \|\phi_i D_{i'j'} \pi_{i'j'}\|_{0,mj} \right) \right. \\
& \left. + \left(F^2_{ij} + F^3_{ij} D_{ij} \right) \|\phi_i\|_{\infty,mj} \right) \|g^*_j\|_{1,mj} h_{mj}. \qquad (5.5)
\end{aligned}$$

As we discussed in Section 4.6, both theorems suggest an adaptive refinement scheme across all meshes used in solving Φ_i and G^*_j, $i = 1, \ldots, N_S$ and $j = 1, \ldots, N_D$. In this respect, we make a modification to the one in Chapter 3 that the new algorithm limits the total number of discretization points in all forward problem meshes, instead of separately limiting the number of discretization points in each mesh used for solving Φ_i or G^*_j. For the adaptive refinement process, the new algorithm is initiated with a set of coarse uniform meshes. With each sweep of refinement and for each source or detector, it computes the error indicator $\varepsilon^i_{MAP,f}(n)$ (or $\varepsilon^j_{MAP,f}(m)$) on every element and computes the average value $\bar{\varepsilon}_{MAP,f}$ of the error indicators on all elements in all meshes. Every element with $\varepsilon^i_{MAP,f}(n) > \bar{\varepsilon}_{MAP,f}$ (or $\varepsilon^j_{MAP,f}(m) > \bar{\varepsilon}_{MAP,f}$) is refined thereafter. By doing so, the resulting meshes provide spatially varying resolution not only within each mesh, but also among all forward problem meshes. The algorithm is stopped when the total number of discretization points in all forward problem meshes reaches a predetermined allowable limit. Algorithm 3 describes the detailed steps of this refinement process in the form of a pseudocode.

Algorithm 3 The pseudocode of the adaptive mesh generation algorithm for the forward problem in the presence of measurement noise.

```
◇ Generate the initial uniform meshes for all forward problems:
$(\Delta^i, N^i_\Delta)$, $\Delta^i = \bigcup_{n=1}^{N^i_\Delta}\{\Delta_n\}$, $i = 1, \ldots, N_S$, and
$(\Delta^{*j}, N^{*j}_\Delta)$, $\Delta^{*j} = \bigcup_{m=1}^{N^{*j}_\Delta}\{\Delta_m\}$, $j = 1, \ldots, N_D$
◇ Set the maximum number of nodes $N^f_{max}$ in all meshes
  while Number of nodes in all meshes less than $N^f_{max}$
    for $i = 1, \ldots, N_S$
      for each element $\Delta_n \in \Delta^i$ with mesh parameter $h_{ni}$
        if first linearization
          . Use analytical solutions for $\phi_i$ and $g^*_j$ and a priori information
          about $\bar{\mu}_{MAP}$ and $\pi_{ij}$ to compute $\varepsilon^i_{MAP,f}(n)$ in (5.4)
        else
          . Use current solution updates $\Phi_i$, $G^*_j$, $\bar{\mu}^D_{MAP}$ and $\Pi_{ij}$ to compute
          $\varepsilon^i_{MAP,f}(n)$ in (5.4)
      end
    end
    for $j = 1, \ldots, N_D$
      for each element $\Delta_m \in \Delta^{*j}$ with mesh parameter $h_{mj}$
        if first linearization
          . Use analytical solutions for $\phi_i$ and $g^*_j$ and a priori information
          about $\bar{\mu}_{MAP}$ and $\pi_{ij}$ to compute $\varepsilon^j_{MAP,f}(m)$ in (5.5)
        else
          . Use current solution updates $\Phi_i$, $G^*_j$, $\bar{\mu}^D_{MAP}$ and $\Pi_{ij}$ to compute
          $\varepsilon^j_{MAP,f}(m)$ in (5.5)
      end
    end
      . Compute $\bar{\varepsilon}_{MAP,f}$
      . Refine the elements with $\varepsilon^i_{MAP,f}(n) > \bar{\varepsilon}_{MAP,f}$ or $\varepsilon^j_{MAP,f}(m) > \bar{\varepsilon}_{MAP,f}$
      . Update the mesh $\Delta^i$, $i = 1, \ldots, N_S$, and $\Delta^{j*}$, $j = 1, \ldots, N_D$
  end
◇ Solve for $\Phi_i$, $i = 1, \ldots, N_S$, and $G^*_j$, $j = 1, \ldots, N_D$
```

To implement this algorithm, we need to make several adjustments: since g_j^*, ϕ_i, $\bar{\mu}_{MAP}$ and π_{ij} in (5.4) and (5.5) also can not be computed exactly, we use analytical solutions for g_j^* and ϕ_i on an unbounded domain (see Appendix F), and the *a priori* information of $\bar{\mu}_{MAP}$ and π_{ij} in the first iteration, and the updated finite-dimensional solutions of them thereafter.

5.3 Adaptive Mesh Generation for Inverse Problem

For the inverse problem, we aim to minimize the M_3 in (4.29), which aries from the inverse problem discretization. For the first adaptive refinement approach, we let $\tilde{\varepsilon}_i$ be the *predetermined* allowable upper bound on $M3$, which represents the allowable error due to the inverse problem discretization. Let

$$\varepsilon_i := \frac{1}{\sqrt{C}N_\Delta}\tilde{\varepsilon}_i. \tag{5.6}$$

Then, we define B_t as follows:

$$B_t := \frac{\varepsilon_i}{\left(\sum_{i,j}^{N_S,N_D} I_{ij}^1 \|G_j^*\Phi_i\|_{0,t} + I_t^2\right)\left(\|\bar{\mu}_{MAP}\|_{1,t} + \sum_{i,j}^{N_S,N_D} \|D_{ij}\pi_{ij}\|_{1,t}\right)}. \tag{5.7}$$

Then for $0 < h_t \leq B_t$, by (4.29) and (5.6), we have

$$M_3 \leq \tilde{\varepsilon}_i.$$

The adaptive mesh refinement procedure of this approach remains the same as the one in Chapter 3.

For the second adaptive refinement approach, the error indicator for each element in the inverse problem mesh is defined as

$$\varepsilon_{MAP,i}(t) := \left(\sum_{i,j}^{N_S,N_D} I_{ij}^1 \|G_j^*\Phi_i\|_{0,t} + I_t^2\right)\left(\|\bar{\mu}_{MAP}\|_{1,t} + \sum_{i,j}^{N_S,N_D} \|D_{ij}\pi_{ij}\|_{1,t}\right) h_t. \tag{5.8}$$

Our new algorithm for the inverse problem again starts from a coarse uniform mesh. In each sweep of refinement, it computes $\varepsilon_{MAP,i}$ for each element and the average value $\bar{\varepsilon}_{MAP,i}$ for all elements, and refines those elements with $\varepsilon_{MAP,i}(t) > \bar{\varepsilon}_{MAP,i}$. The algorithm stops when the total number of discretization points exceeds the predetermined allowable limit. Algorithm 4 describes the detailed steps of this refinement process in the form of a pseudocode. Finally, we note that we use the same strategy as that we used in the previous section to approximate $\bar{\mu}_{MAP}$ and π_{ij} in calculating B_t or $\varepsilon_{MAP,i}$.

Algorithm 4 The pseudocode of the adaptive mesh generation algorithm for the inverse problem in the presence of measurement noise.

```
◇ Generate an initial uniform mesh: $(\Delta, N_\Delta)$, $\Delta = \bigcup_{t=1}^{N_\Delta}\{\Delta_t\}$
◇ Set the maximum number of nodes $N^i_{max}$
  while Number of nodes $N$ less than $N^i_{max}$
    for each element $\Delta_t \in \Delta$ with mesh size parameter $h_t$
      if first linearization
        . Use current solution updates $\Phi_i$ and $G^*_j$, and a priori information
          about $\pi_{ij}$ and $\bar{\mu}_{MAP}$ to compute $\varepsilon_{MAP,i}(t)$ in (5.8 )
      else
        . Use current solution updates $\Phi_i$, $G^*_j$, $\bar{\mu}^D_{MAP}$ and $\Pi_{ij}$ to compute
          $\varepsilon_{MAP,i}(t)$ in (5.8 )
    end
      . Compute $\bar{\varepsilon}_{MAP,i}$
      . Refine the elements with $\varepsilon_{MAP,i}(t) > \bar{\varepsilon}_{MAP,i}$
      . Update the mesh $\Delta$
  end
◇ Solve for $\mu^D_{MAP}$ and $\Pi_{ij}$
```

5.4 Computational Complexity

In this section, we briefly discuss the computational complexity of the adaptive mesh generation algorithms described in the previous sections. In both Algorithms 3 and 4, one needs to solve π_{ij} for $i = 1, 2, \ldots, N_S$ and $j = 1, 2, \ldots, N_D$. These π_{ij}s can be computed once numerically given the source-detector geometry and background optical properties, and stored for the computation of the error indicators in adaptive mesh refinement. To further improve the computational efficiency of Algorithms 3 and 4, we consider the following approximations to (5.4), (5.5) and (5.8):

$$\begin{aligned}
\varepsilon^{i}_{MAP,f}(n) &\approx \sum_{j=1}^{N_D} \left(F^1_{ij} \left(\|g_j^* \bar{\mu}_{MAP}\|_{0,ni} + \left\| g_j^* \sum_{i',j'}^{N_S,N_D} D_{i'j'}\pi_{i'j'} \right\|_{0,ni} \right) \right. \\
&\quad \left. + \left(F^2_{ij} + F^3_{ij} D_{ij}\right) \|g_j^*\|_{\infty,ni} \right) \|\phi_i\|_{1,ni} h_{ni} \qquad (5.9) \\
\varepsilon^{j}_{MAP,f}(m) &\approx \sum_{i=1}^{N_S} \left(F^1_{ij} \left(\|\phi_i \bar{\mu}_{MAP}\|_{0,mj} + \left\| \phi_i \sum_{i',j'}^{N_S,N_D} D_{i'j'}\pi_{i'j'} \right\|_{0,mj} \right) \right. \\
&\quad \left. + \left(F^2_{ij} + F^3_{ij} D_{ij}\right) \|\phi_i\|_{\infty,mj} \right) \|g_j^*\|_{1,mj} h_{mj}, \qquad (5.10) \\
\varepsilon_{MAP,i}(t) &\approx \left(\sum_{i,j}^{N_S,N_D} I^1_{ij} \|G_j^* \Phi_i\|_{0,t} + I^2_t \right) \left(\|\bar{\mu}_{MAP}\|_{1,t} + \left\| \sum_{i,j}^{N_S,N_D} D_{ij}\pi_{ij} \right\|_{1,t} \right) h_t. \qquad (5.11)
\end{aligned}$$

It is clear that $\sum_{i,j}^{N_S,N_D} D_{ij}\pi_{ij}$ can be solved by using

$$\mathcal{F}_{MAP}\left(\psi, \sum_{i,j}^{N_S,N_D} D_{ij}\pi_{ij}\right) = \left(\psi, \mathcal{A}^* \mathbf{\Sigma}_{\boldsymbol{\varepsilon}}^{-1} \sum_{i,j}^{N_S,N_D} D_{ij}\mathbf{e}_{ij}\right).$$

In this case, using (5.9), (5.10) and (5.11) as the error indicators in Algorithms 3 and 4 only requires one image reconstruction procedure to solve $\sum_{i,j}^{N_S,N_D} D_{ij}\pi_{ij}$, instead of solving each π_{ij} separately.

Based on the modified error indicators (5.9), (5.10) and (5.11), we analyze the computational complexity of our adaptive mesh generation algorithms. First of all, in (5.9) and (5.10), $\|g_j^*\phi_i\|_0$, $i = 1, \dots, N_S$ and $j = 1, \dots, N_D$, can be computed before the mesh refinement. Therefore, we still have $\mathcal{O}(N^i_\Delta N_D)$ and $\mathcal{O}(N^{*j}_\Delta N_S)$ complexity for adaptively refining the mesh based on (5.9) and (5.10) in each refinement sweep. We can further reduce them to $\mathcal{O}(N^i_\Delta)$ and $\mathcal{O}(N^{*j}_\Delta)$ by using the approximations to (5.9) and (5.10) suggested in Section 3.6:

$$\begin{aligned}\varepsilon^{i}_{MAP,f}(n) &\approx \left(\left(\left\|\sum_{j=1}^{N_D} F^1_{ij} g^*_j \bar{\mu}_{MAP}\right\|_{0,ni} + \left\|\sum_{j=1}^{N_D} F^1_{ij} g^*_j \sum_{i',j'}^{N_S,N_D} D_{i'j'}\pi_{i'j'}\right\|_{0,ni}\right)\right. \\ &\quad \left. + \sum_{j=1}^{N_D}\left(F^2_{ij} + F^3_{ij} D_{ij}\right)\|g^*_j\|_{\infty,ni}\right), \|\phi_i\|_{1,ni} h_{ni} \qquad (5.12) \\ \varepsilon^{j}_{MAP,f}(m) &\approx \left(\left(\left\|\sum_{i=1}^{N_S} F^1_{ij} \phi_i \bar{\mu}_{MAP}\right\|_{0,mj} + \left\|\sum_{i=1}^{N_S} F^1_{ij} \phi_i \sum_{i',j'}^{N_S,N_D} D_{i'j'}\pi_{i'j'}\right\|_{0,mj}\right)\right. \\ &\quad \left. + \sum_{i=1}^{N_S}\left(F^2_{ij} + F^3_{ij} D_{ij}\right)\|\phi_i\|_{\infty,mj}\right) \|g^*_j\|_{1,mj} h_{mj}, \qquad (5.13)\end{aligned}$$

Second, in (5.11), $\|G^*_j \Phi_i\|_0$, $i = 1, \ldots, N_S$ and $j = 1, \ldots, N_D$, can also be computed before the mesh refinement. Therefore, the computational complexity for refining the inverse problem mesh using (5.11) in each sweep of the mesh refinement is $\mathcal{O}(N_\Delta N_S N_D)$. It can also be reduced to $\mathcal{O}(N_\Delta)$ by a further approximation of (5.11):

$$\varepsilon_{MAP,i}(t) \approx \left(\left\|\sum_{i,j}^{N_S,N_D} I^1_{ij} G^*_j \Phi_i\right\|_{0,t} + I^2_t\right)\left(\|\bar{\mu}_{MAP}\|_{1,t} + \left\|\sum_{i,j}^{N_S,N_D} D_{ij}\pi_{ij}\right\|_{1,t}\right) h_t. \qquad (5.14)$$

Therefore, using the approximated error indicators in (5.12), (5.13) and (5.14), our new adaptive mesh generation algorithms have the same computational complexity as those of the algorithms in Chapter 3 and the conventional adaptive meshing method.

5.5 Numerical Simulation

To demonstrate the improvements in the FDOT reconstruction accuracy by using the adaptive mesh generation algorithms proposed in this chapter in the presence of measurement noise, we performed a set of numerical simulation studies. Specifically, we simulated a realistic noise situation in FDOT imaging, and evaluated the

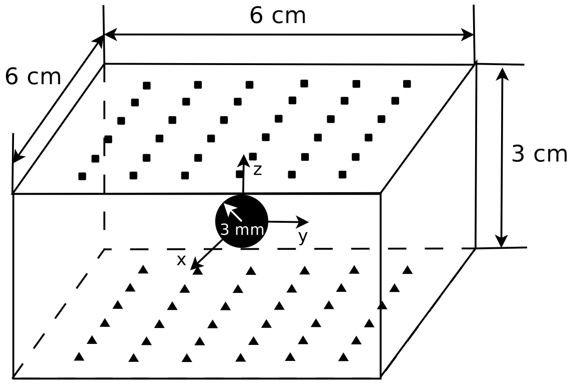


Figure 5.1: The simulation setups used for Simulation Studies 3 and 4. The squares and triangles denote the detectors and sources, respectively.

accuracy of FDOT reconstruction by calculating the MSE in the reconstructed images due to discretization. In the simulation study, we compared our adaptive mesh generation algorithms proposed in this chapter with the uniform meshing scheme and the one developed in Chapter 3, to show the practical advantages of the algorithms in this chapter in the presence measurement noise.

5.5.1 Simulation Setup

In the numerical simulation, we considered a 6 cm × 6 cm × 3 cm cubic domain Ω shown in Figure 5.1. We set the homogeneous background absorption coefficient $\mu_{axe} = \mu_{ame} = 0.05$ cm^{-1} and diffusion coefficient $D = 0.0410$ for both excitation and emission wavelengths, and set the refractive index mismatch parameter $\rho = 3$ for the boundary $\partial\Omega$. At the center of the domain, we placed a fluorophore heterogeneity with 3 mm radius, constant absorption coefficient $\mu_{axf} = 0.015$ cm^{-1} and quantum efficiency $\eta = 0.05$. In the rest of the domain, we assumed $\mu_{axf} = 0$. To reconstruct the fluorophore concentration image, we placed 36 sources and 36 detectors evenly on two 6 × 6 grids at the bottom and top surfaces of the domain, as shown in Figure 5.1. We simulated both excitation and emission light fields by solving the coupled diffusion equations (2.5) and (2.6) with their corresponding boundary conditions (2.7) and (2.8), using the parameters above on a fine uniform grid with 81×81×41 nodes.

To simulate measurement noise, we considered the noise model (4.30) in Section 4.7. We further define the SNR of the measurements as

$$\mathrm{SNR}_{ij} = 10 \log_{10} \frac{|\Gamma_{0,ij}|^2}{\sigma^2_{\boldsymbol{\varepsilon},ij}} = 10 \log_{10} \frac{|\Gamma_{0,ij}|}{\alpha}. \tag{5.15}$$

Note that, each measurement, Γ_{ij}, has a different SNR proportional to $\log_{10} |\Gamma_{0,ij}|$.

In Simulation Study 3, we simulated the noise $\boldsymbol{\varepsilon}$ with 3 different values of α: 5×10^{-11}, 1×10^{-9} and 5×10^{-9}, corresponding to approximately 40, 26 and 20 dB average SNR over all measurements Γ_{ij}, $i = 1, \ldots, N_S$ and $j = 1, \ldots, N_D$. For each value of α, we generated 100 different realizations of noise and obtain three sets of noise contaminated measurements with approximately 1%, 5% and 10% noise. By changing the noise level, we aimed to show the effect of measurement noise on the adaptive mesh generation as well as the image reconstruction. In the FDOT reconstruction, we reconstructed the optical image using the MAP estimator of fluorophore concentration and considered a simplified *a priori* model for the fluorophore concentration:

$$\kappa(\boldsymbol{r}) = \kappa_0,$$

where $\kappa_0 = 5 \times 10^{-6}$ is a constant chosen empirically.

In Simulation Study 4, we simulated the noise $\boldsymbol{\varepsilon}$ with $\alpha = 5 \times 10^{-11}$. Instead of choosing a deterministic value for the absorption coefficient of the fluorophore, μ_{axf}, we randomly chose its value from a Gaussian distribution, $\mathcal{N}(\mu^0_{axf}, \sigma^2_{axf})$, and dropped the negative values generated. We set $(\mu^0_{axf}, \sigma_{axf})$ to three different values: (0.005 cm^{-1}, 0.002 cm^{-1}), (0.015 cm^{-1}, 0.006 cm^{-1}), and (0.025 cm^{-1}, 0.010 cm^{-1}), to simulate three cases of different fluorophore variance. For each case, we generated 100 realizations of μ_{axf}.

We further considered the following *a priori* model for the fluorophore concentration:

$$\kappa(\boldsymbol{r}) = \begin{cases} \kappa_f & \boldsymbol{r} \in \Omega_f \\ \kappa_b & \boldsymbol{r} \in \Omega_b \end{cases},$$

where κ_f and κ_b are constant, $\Omega_f \cup \Omega_b = \Omega$, $\Omega_f \cap \Omega_b = \emptyset$, Ω_f denotes the fluorophore region, and Ω_b denotes the background region. We chose $\kappa_b = \kappa_0 = 5 \times 10^{-6}$, and chose κ_f with 3 different values: 5×10^{-6}, 3.6×10^{-5}, and 1×10^{-4}. Note that in the latter two cases of Simulation Study 4, we have the specific *a priori* information on fluorophore concentration that the variance of the fluorophore concentration inside the heterogeneity region, Ω_f, is higher than that of the background region, Ω_b. By changing κ_f, we aimed to show the effect of *a priori* information of fluorophore concentration on the adaptive mesh generation as well as the image reconstruction.

Finally, we note that we again performed our simulation study using deal.II FEM C++ library [98] and used hexahedral finite elements with trilinear Lagrange basis functions to discretize both forward and inverse problems.

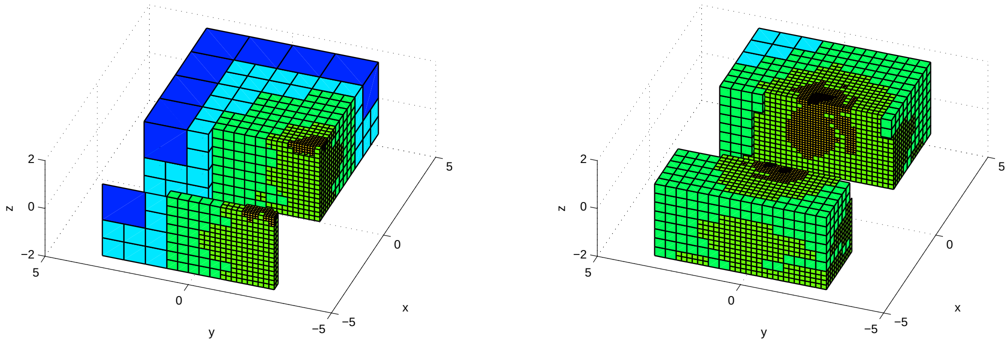
5.5.2 Mesh Generation

We used three different types of coarse meshes: uniform meshes, the adaptive meshes generated by the algorithms in Chapter 3, and the adaptive meshes generated by the algorithms in this chapter, to discretize the forward and inverse problems. For the forward problem, the total number of nodes in the meshes used in solving all Φ_i and G_j^*, $i = 1, \ldots, N_S$ and $j = 1, \ldots, N_D$, ranged from 500,000 to 650,000 (roughly 7,000 to 9,000 for each mesh); and for the inverse problem, it ranged from 2,000 to 3,000. Note that the uniform meshes used for solving the forward and inverse problems have 25×25×13 nodes and 17×17×9 nodes, respectively. For performance evaluation, we considered three different image reconstruction scenarios corresponding to these three different meshing schemes:

1. We used the coarse uniform mesh shown in Figure 3.3(a) to discretize the forward problem and the coarse uniform mesh shown in Figure 3.3(b) to discretize the inverse problem. We denote the resulting image by $\mu^D_{MAP,U}$.

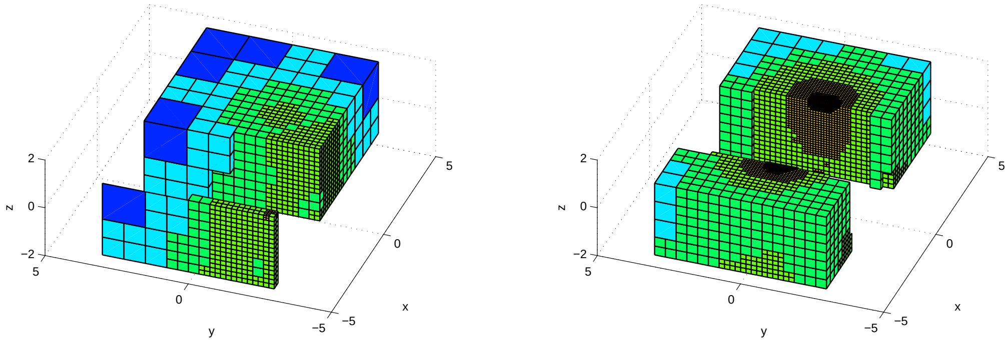
2. We used the adaptive meshes generated by Algorithm 3 (see Figures 5.2(a)-5.2(d)) to discretize the forward problem and the adaptive mesh generated by Algorithm 4 (see Figures 5.5(a), 5.5(b) and 5.5(d)) to discretize the inverse problem. We denote the resulting image by $\mu^D_{MAP,NA}$.

3. We used the adaptive meshes generated by Algorithm 1 (see Figures 5.2(e) and 5.2(f)) to discretize the forward problem and the adaptive mesh generated



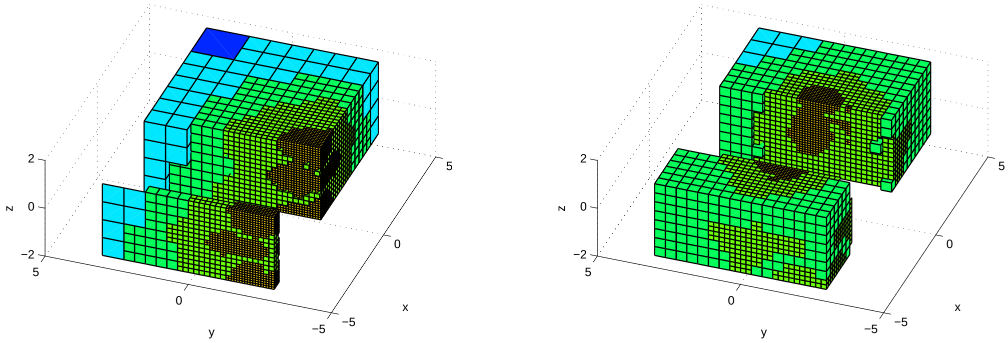
(a) The adaptive mesh with 3,289 nodes generated by our algorithm in this chapter for the detector located at (-2.5,-2.5,1.5) for the 1% noise case.

(b) The adaptive mesh with 18,876 nodes generated by our algorithm in this chapter for the detector located at (-0.5,-0.5,1.5) for the 1% noise case.



(c) The adaptive mesh with 3,588 nodes generated by our algorithm in this chapter for the detector located at (-2.5,-2.5,1.5) for the 10% noise case.

(d) The adaptive mesh with 18,054 nodes generated by our algorithm in this chapter for the detector located at (-0.5,-0.5,1.5) for the 10% noise case.



(e) The adaptive mesh with 8,304 nodes generated by the algorithm in Chapter 3 for the detector located at (-2.5,-2.5,1.5).

(f) The adaptive mesh with 7,973 nodes generated by the algorithm in Chapter 3 for the detector located at (-0.5,-0.5,1.5).

Figure 5.2: Examples of the adaptive meshes for the forward problem used in Simulation Study 3. The mesh is cut through to show the mesh structure inside.

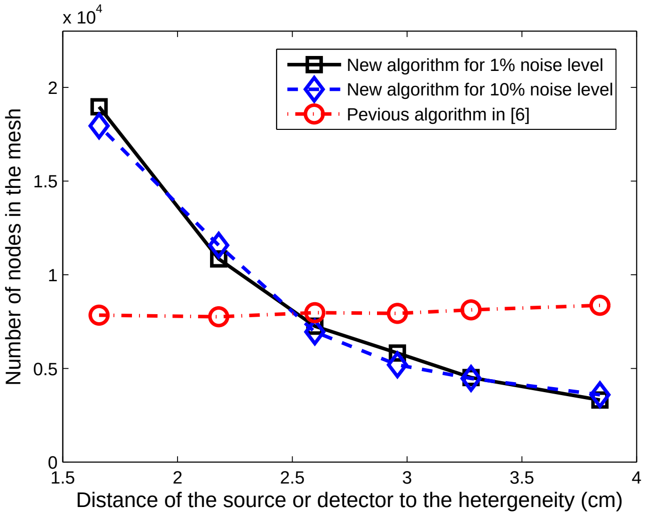


Figure 5.3: The relationship between the number of nodes in the forward adaptive mesh for a certain source (or detector) and the distance of the source (or detector) to the center of the fluorophore heterogeneity in the 1% and 10% noise cases in Simulation Study 3.

by Algorithm 2 (see Figures 5.5(c) and 5.5(e)) to discretize the inverse problem. We denote the resulting image by $\mu^D_{MAP,A}$.

For the forward problem, the examples of the adaptive meshes generated for the detectors located at $(-2.5, -2.5, 1.5)$ and $(-0.5, -0.5, 1.5)$ in the 1% and 10% noise cases of Simulation Study 3 are shown in Figure 5.2. Figures 5.2(a) - 5.2(d) show the meshes generated by the algorithm in this chapter. We observe there are more nodes in the meshes for the detector located at $(-2.5, -2.5, 1.5)$ than in the meshes for the detector located at $(-0.5, -0.5, 1.5)$. Figures 5.2(e) and 5.2(f) show the corresponding meshes generated by the algorithm in Chapter 3, and these two meshes have approximately same number of nodes.

We plotted the relationship between the numbers of nodes in the meshes generated by the algorithms in this chapter and Chapter 3 for a certain source (or detector) and the distance of that source (or detector) to the center of the fluorophore heterogeneity in Figure 5.3. Note that for the sources and detectors which have the same distance to the heterogeneity, we plotted the average number of nodes in the corresponding meshes. For the algorithm in this chapter, we observe that the closer the sources or detectors to the heterogeneity, the larger the number of nodes

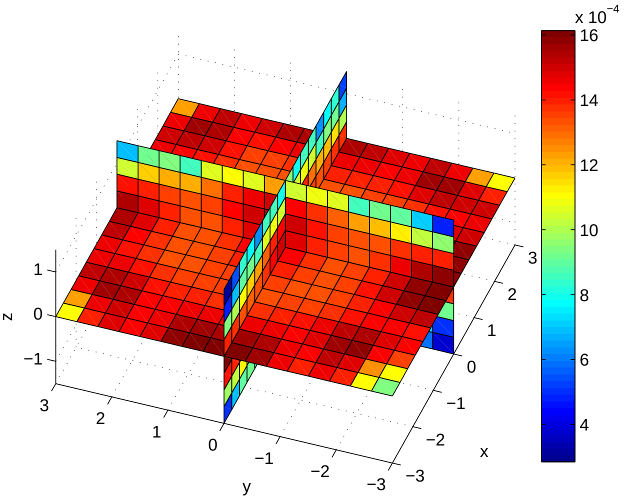


Figure 5.4: The cross-sections of the image, $\sum_{i,j}^{N_S,N_D} D_{ij}\pi_{ij}$, reconstructed using coarse uniform meshes in the 10% noise case in Simulation Study 3.

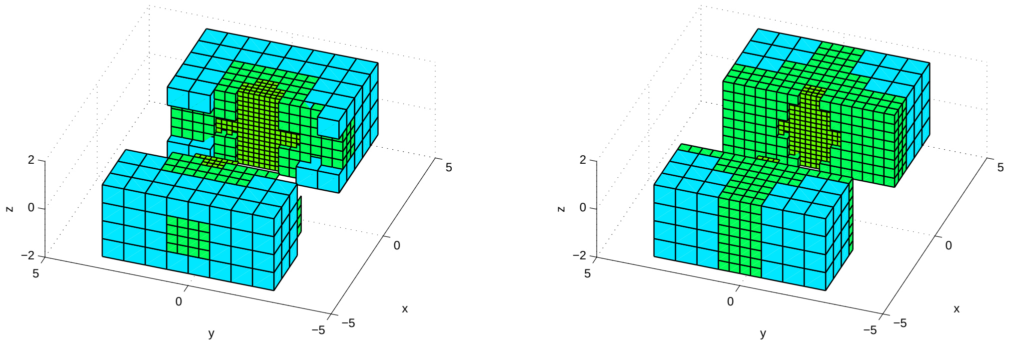
is in the associated meshes. This can be explained with the fact that for those source-detector pairs closer to the heterogeneity, the measurements have higher SNR. As a result, the algorithm in this chapter generates finer meshes for these source-detector pairs, so that the accuracies of the corresponding forward problem solutions can match the accuracy of the measurements. This results in the forward problem meshes generated by the algorithm in this chapter with varying resolution for different source-detector pairs. We note that the similar results of the adaptive mesh generation for the forward problem were also observed in Simulation Study 4.

In Simulation Study 3, we further observe that the algorithm in Chapter 3 generates the same mesh for different noise levels as shown in Figures 5.2(e) - 5.2(f). This indicates that, when the noise level changes from 1% to 10%, the change in the reconstructed image does not have enough impact on the adaptive mesh refinement to change the mesh finally generated. This can be further validated by the image reconstruction results shown in Figures 5.6 and 5.7 for 1% and 10% measurement noise level cases. In these reconstruction results, we observe that the major difference in the reconstructed images is the change in the background variability when the noise level changes. On the other hand, in Figures 5.2(a) - 5.2(d), we note the difference between the adaptive meshes generated by the algorithm in this chapter for different noise levels, which illustrates the impact of noise level on the meshes generated by this new algorithm. In this algorithm, since the value of the coefficient D_{ij} increases as $\sigma^2_{\mathcal{E},ij}$ increases, the images $D_{ij}\pi_{ij}$, $i = 1, \ldots, N_S$ and $j = 1, \ldots, N_D$, have more contribution to the mesh refinement when the noise level is high. Figure 5.4 shows the cross-section of $\sum_{i,j} D_{ij}\pi_{ij}$ for the 10% noise case. These results indicate that the algorithm in this chapter can adaptively refine the meshes according to the measurement noise level. In Simulation Study 4, as we changed the fluorophore variance, we also observed a change in adaptive mesh generation similar to that in Simulation Study 3.

For the inverse problem, we show the examples of the adaptive meshes in Figure 5.5. Figures 5.5(a) and 5.5(b) show two different meshes generated by the algorithm in this chapter for the 1% and 10% noise level cases in Simulation Study 3, respectively. Similarly, we note that the difference in these two meshes is mainly due to the change of noise level. The algorithm in Chapter 3 does not take into account noise and generates the same mesh for different noise levels as shown in Figure 5.5(c).

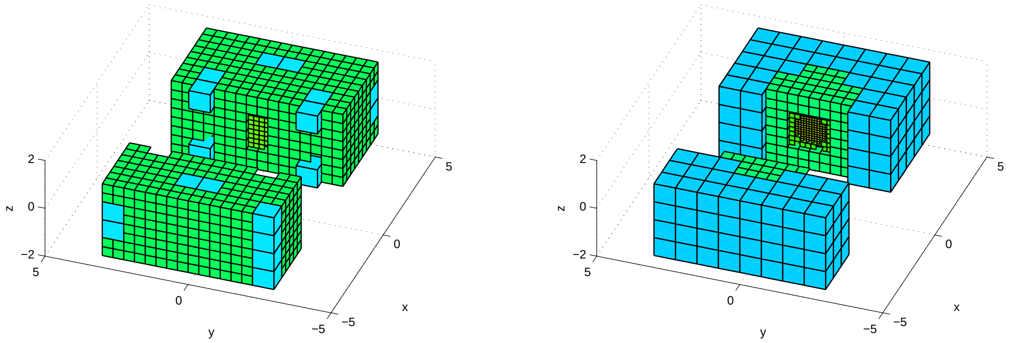
Figures 5.5(e) and 5.5(d) show the meshes generated by algorithms in Chapters 3 and 5 when the specific *a priori* information was used in Simulation Study 4. Comparing these two meshes with the meshes generated in Simulation Study 3, we observe that both algorithms generated different meshes for the different *a priori* information. However, we note that the difference in the meshes generated by the algorithm in Chapter 3 is only due to the dramatic change in the reconstructed image (see Figure 5.8) resulting from the specific *a priori* information used in FDOT reconstruction. On the other hand, the difference in the meshes generated by the algorithm in this chapter is due to changes in both the reconstructed image and *a priori* information. To conclude, the results in Figure 5.5 indicate that the algorithm in this chapter takes into account the noise statistics and *a priori* information of fluorophore concentration, and can adaptively refine the mesh according to these two factors.

In the following sections, we present the simulation results and demonstrate the improvements in the reconstruction accuracy as well as the convergence rate of



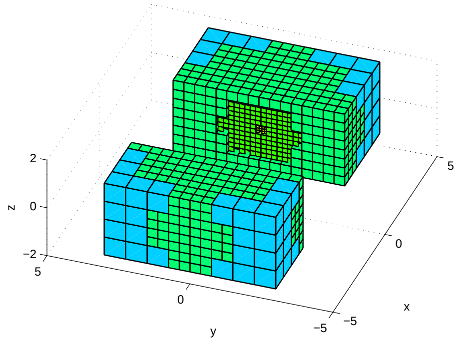
(a) The adaptive mesh with 2,721 nodes generated by the algorithm in Chapter 5 for the inverse problem for the 1% noise case in Simulation Study 3.

(b) The adaptive mesh with 2,785 nodes generated by the algorithm in Chapter 5 for the inverse problem for the 10% noise case in Simulation Study 3.



(c) The adaptive mesh with 2,652 nodes generated by the algorithm in Chapter 3 for the inverse problem in Simulation Study 3.

(d) The adaptive mesh with 2,836 nodes generated by the algorithm in Chapter 3 for the inverse problem with $\mu^0_{axf} = 0.025$ cm^{-1} and $\sigma_{axf} = 0.01$ cm^{-1} in Simulation Study 4.



(e) The adaptive mesh with 2,991 nodes generated by the algorithm in Chapter 5 for the inverse problem with $\mu^0_{axf} = 0.025$ cm^{-1} and $\sigma_{axf} = 0.01$ cm^{-1} in Simulation Study 4.

Figure 5.5: Examples of the adaptive meshes for the inverse problem used in Simulation Studies 3 and 4. The mesh is cut through to show the mesh structure inside.

Table 5.1: MSE, bias and variance of the images reconstructed by using different meshes in Simulation Study 3.

Noise Level	Images	Bias^2 ($\times 10^{-4}$)		Var ($\times 10^{-6}$)		MSE ($\times 10^{-4}$)	
1%	$\mu^D_{MAP,U}$	3.295	100%	0.037	100%	3.295	100%
	$\mu^D_{MAP,NA}$	0.857	26.00%	0.009	23.49%	0.857	26.00%
	$\mu^D_{MAP,A}$	2.040	61.93%	0.038	101.45%	2.041	61.93%
5%	$\mu^D_{MAP,U}$	3.251	100%	0.893	100%	3.260	100%
	$\mu^D_{MAP,NA}$	0.753	23.16%	0.199	22.27%	0.755	23.16%
	$\mu^D_{MAP,A}$	2.028	62.38%	0.990	110.92%	2.038	62.52%
10%	$\mu^D_{MAP,U}$	3.217	100%	3.670	100%	3.253	100%
	$\mu^D_{MAP,NA}$	0.768	23.89%	0.872	23.77%	0.777	23.89%
	$\mu^D_{MAP,A}$	2.045	63.57%	4.301	117.19%	2.088	64.18%

Table 5.2: MSE, bias and variance of the images reconstructed by using different meshes in Simulation Study 4.

$(\mu^0_{axf}, \sigma_{axf})$	Images	Bias^2 ($\times 10^{-4}$)		Var ($\times 10^{-4}$)		MSE ($\times 10^{-4}$)	
(0.005,0.002)	$\mu^D_{MAP,U}$	1.406	100%	0.270	100%	1.676	100%
	$\mu^D_{MAP,NA}$	0.319	22.70%	0.060	22.37%	0.380	22.64%
	$\mu^D_{MAP,A}$	0.860	61.15%	0.263	97.30%	1.122	66.97%
(0.015,0.006)	$\mu^D_{MAP,U}$	3.508	100%	0.638	100%	4.146	100%
	$\mu^D_{MAP,NA}$	0.953	27.15%	0.163	25.49%	1.115	26.90%
	$\mu^D_{MAP,A}$	2.330	66.41%	0.700	109.74%	3.030	73.08%
(0.025,0.010)	$\mu^D_{MAP,U}$	4.613	100%	0.801	100%	5.413	100%
	$\mu^D_{MAP,NA}$	1.396	30.26%	0.224	28.01%	1.620	29.92%
	$\mu^D_{MAP,A}$	3.205	69.49%	0.872	108.93%	4.078	75.32%

the discretized inverse problem solution to the exact inverse problem solution, when the adaptive meshes refined by taking into account the measurement noise as well as *a priori* information on fluorophore concentration are used in FDOT reconstruction.

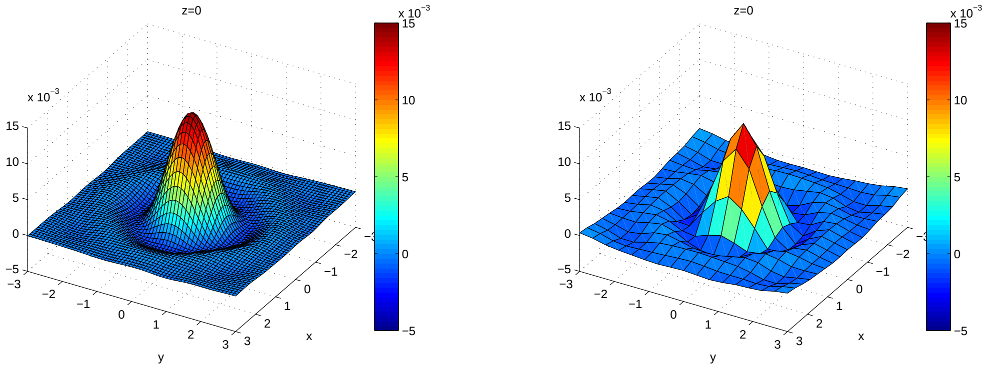
5.5.3 Simulation Results

5.5.3.1 Reconstruction Accuracy

To obtain the exact solutions of the forward and inverse problems, we solved the forward and inverse problems on a fine mesh with 61 × 61 × 31 nodes. We assumed that the error due to discretization in the resulting image, denoted by μ_{MAP} is negligible with respect to the images reconstructed using the three types of coarse meshes; and used this image as a baseline to compute the MSE due to discretization. In both Simulation Studies 3 and 4, we calculated the bias, variance and the MSE of the reconstructed images for each set of reconstructions by averaging all reconstructed image samples for 100 realizations of noise and fluorophore concentration value. The results are tabulated in Tables 5.1 and 5.2. Additionally, we tabulated the percentage of each quantity as compared to $\mu^D_{MAP,U}$, the image reconstructed by using the coarse uniform meshes: The left column is the absolute value, and the right column is the corresponding percentage. In Simulation Study 3, the results in Table 5.1 show that the bias squares of the images, reconstructed using different types of meshes, remain at a fixed level when the noise level changes, while the variances of the images increase as the noise level increases. In Simulation Study 4, the results in Table 5.2 show that both bias squares and variances of the reconstructed images increase as the fluorophore variance increases. The results in both Tables 5.1 and 5.2 show that the bias square, variance, as well as the MSE of $\mu^D_{MAP,NA}$ are approximately reduced by 70% to 75% as compared to $\mu^D_{MAP,U}$, when the algorithms in this chapter are used. On the other hand, the algorithms in Chapter 3 provide approximately 30% to 40% reduction in the bias square, but no reduction in the variance of $\mu^D_{MAP,A}$ with respect to $\mu^D_{MAP,U}$.

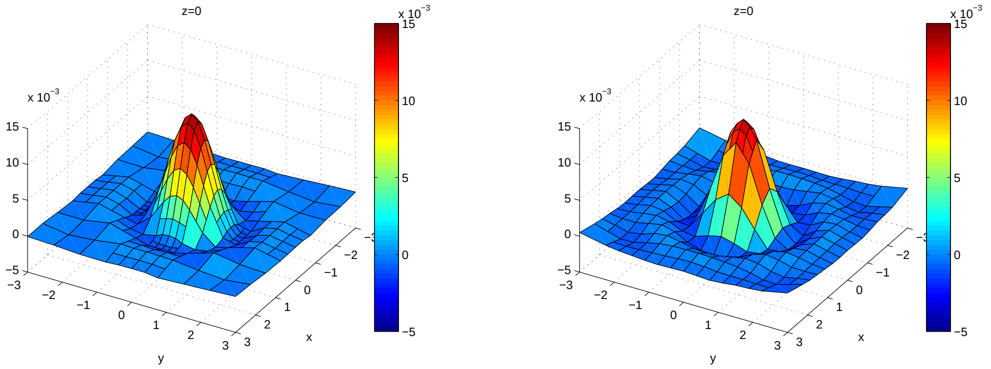
Figures 5.6 and 5.7 show the cross-section of the sample images at $z = 0$ plane reconstructed by using different types of meshes when the noise level is 1% and 10% in Simulation Study 3. The cross-section of the baseline images are shown in Figures 5.6(a) and 5.7(a). We observe that the variability of images in Figure 5.7 is more visible as compared to that of the images in Figure 5.6 due to increased noise level in the measurements. The shape of the small fluorophore heterogeneity is better resolved in $\mu^D_{MAP,A}$ and $\mu^D_{MAP,NA}$ as compared to the one in $\mu^D_{MAP,U}$, due to the spatially varying resolution provided by the adaptive meshes. Additionally, we observe a higher background variability in $\mu^D_{MAP,A}$ than that of $\mu^D_{MAP,NA}$ in Figures 5.7(d) and 5.7(c), while the difference between the images in Figures 5.6(d) and 5.6(c) is not as noticeable due to low noise level.

Figure 5.8 shows the cross-section of the sample images at $z = 0$ plane reconstructed by using different types of meshes when $\mu_{axf} = 0.033$ cm^{-1} in Simulation



(a) The baseline image.

(b) The image reconstructed using the uniform meshes.

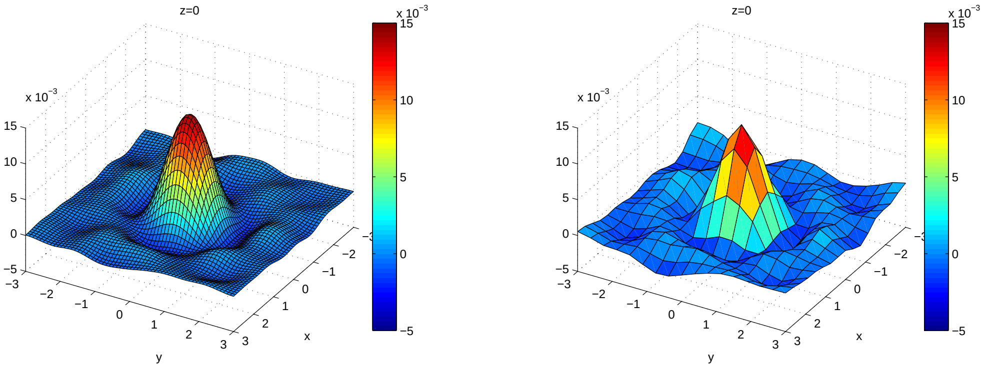


(c) The image reconstructed using the adaptive meshes generated by the algorithms in Chapter 5.

(d) The image reconstructed using the adaptive meshes generated by the algorithms in Chapter 3.

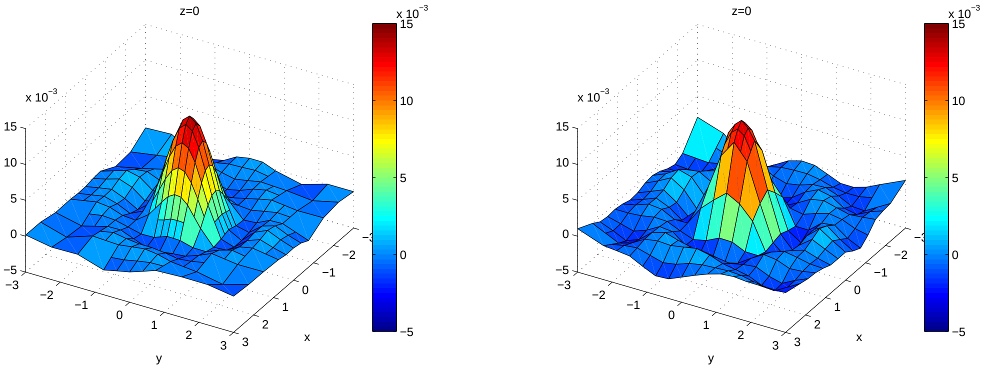
Figure 5.6: The reconstruction results of Simulation Study 3, Case 1, on plane $z = 0$, at 1% noise level.

Study 4. Since the specific *a priori* information on fluorophore concentration is used in the image reconstruction, all the reconstructed images, shown in Figure 5.8, are reconstructed with a narrower peak at the fluorophore heterogeneity region as compared to the images shown in Figures 5.6 and 5.7. We observe that the fluorophore heterogeneity in $\mu^D_{MAP,U}$ is reconstructed in a pyramid shape due to the fact that the size of fluorophore heterogeneity is smaller than the size of element in the uniform mesh used to discretize the inverse problem. On the other hand, because of the spatially varying resolution provided by the adaptive meshes, the fluorophore heterogeneities in $\mu^D_{MAP,NA}$ and $\mu^D_{MAP,A}$ are better revolved than that in $\mu^D_{MAP,U}$. Furthermore, $\mu^D_{MAP,NA}$ reconstructed by using the algorithms in this chapter is closer to the baseline image than $\mu^D_{MAP,A}$ reconstructed by the algorithms



(a) The baseline image.

(b) The image reconstructed using the uniform meshes.



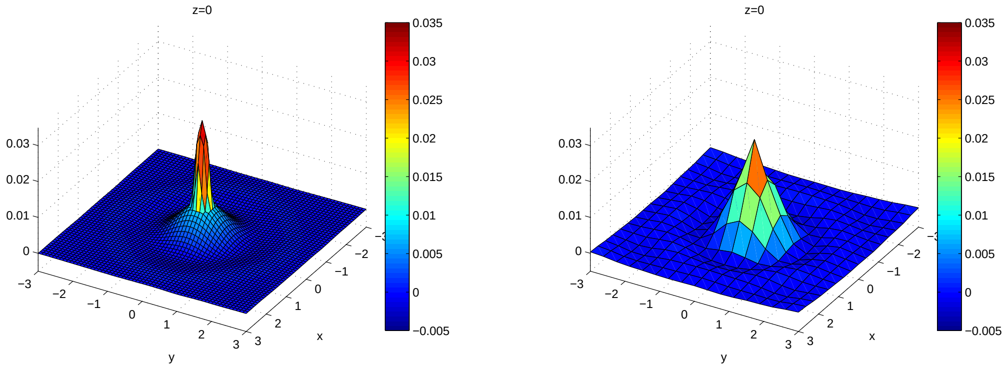
(c) The image reconstructed using the adaptive meshes generated by the algorithms in Chapter 5.

(d) The image reconstructed using the adaptive meshes generated by the algorithms in Chapter 3.

Figure 5.7: The reconstruction results of Simulation Study 3, Case 3, on plane $z = 0$, at 10% noise level.

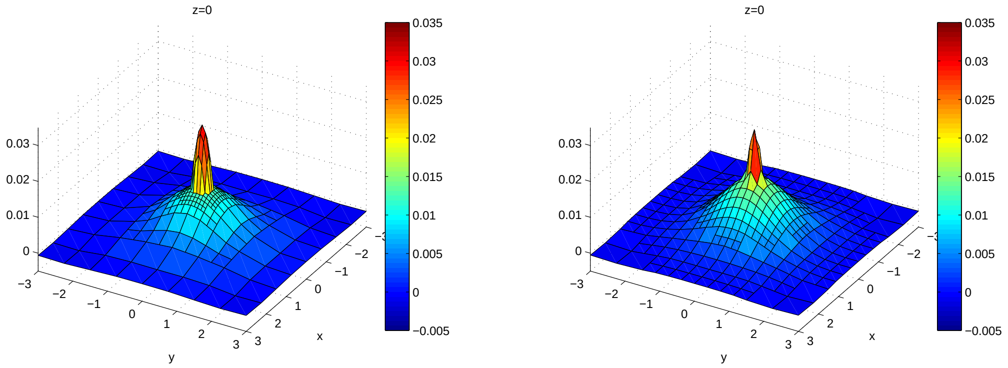
in Chapter 3.

The reconstruction results in Figures 5.6, 5.7, and 5.8 can be seen more clearly in Figures 5.9 and 5.10, where the reconstructed images along the y-axis on $z = 0$ plane are shown. The solid lines in Figure 5.9 and 5.10 represent the baseline image μ_{MAP} which is assumed to have negligible error due to discretization. We observe that the image, $\mu^D_{MAP,NA}$, is the best approximation to μ_{MAP} in all three reconstructed images, which has higher response at the center of the fluorophore heterogeneity and lower background variance, as compared to those of $\mu^D_{MAP,U}$ and $\mu^D_{MAP,A}$.



(a) The baseline image.

(b) The image reconstructed using the uniform meshes.



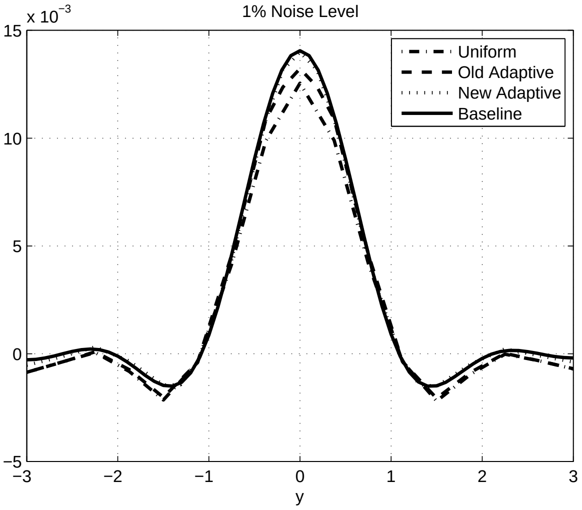
(c) The image reconstructed using the adaptive meshes generated by the algorithms in Chapter 5.

(d) The image reconstructed using the adaptive meshes generated by the algorithms in Chapter 3.

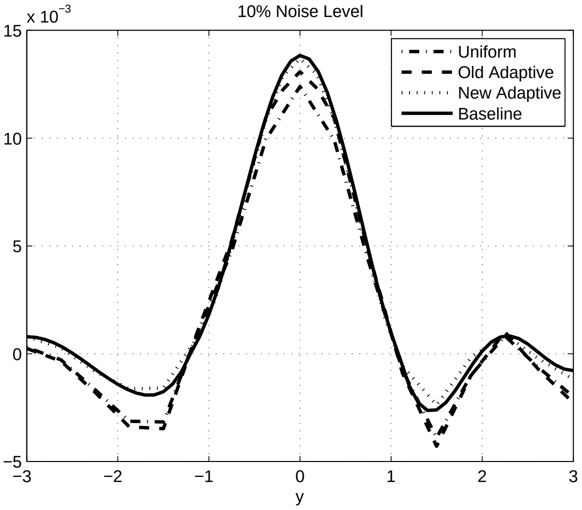
Figure 5.8: The reconstruction results of Simulation Study 4, Case 3, on plane $z = 0$, with $\mu_{axf} = 0.033$ cm^{-1} and $\kappa_f = 1 \times 10^{-4}$.

5.5.3.2 Convergence Rate

To study the convergence of the discretized inverse problem solution to the exact solution when different types of meshes are used in FDOT reconstruction in the presence of noise, we solved both forward and inverse problems on a series of meshes with gradually increasing discretization levels in each reconstruction scenario. Note that, for each discretization level, the mesh for the forward problem is always slightly finer than the mesh for the inverse problem. At the same time, we calculated the MSE between the two images reconstructed on two consecutive discretization levels, and normalized it with the mean $L^2(\Omega)$ norm square of the image reconstructed on the lower discretization level. We, again, refer to this as the percentage error between the mesh refinement. This percentage error decreases as the discretization



(a) The images along the y-axis reconstructed using measurements with 1% noise level in Simulation 3, Case 1, at 1% noise level.



(b) The images along the y-axis reconstructed using measurements with 10% noise level in Simulation 3, Case 3.

Figure 5.9: The profiles of the reconstructed images μ_{MAP}, $\mu^D_{MAP,U}$, $\mu^D_{MAP,A}$ and $\mu^D_{MAP,NA}$ along the y-axis in Simulation Study 3, Case 1 and Case 3.

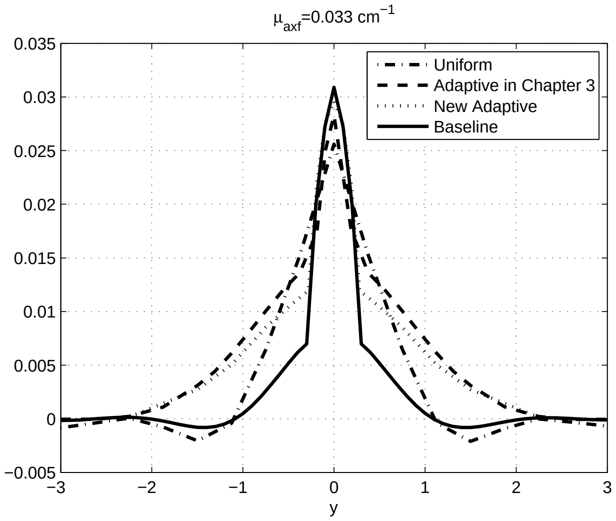
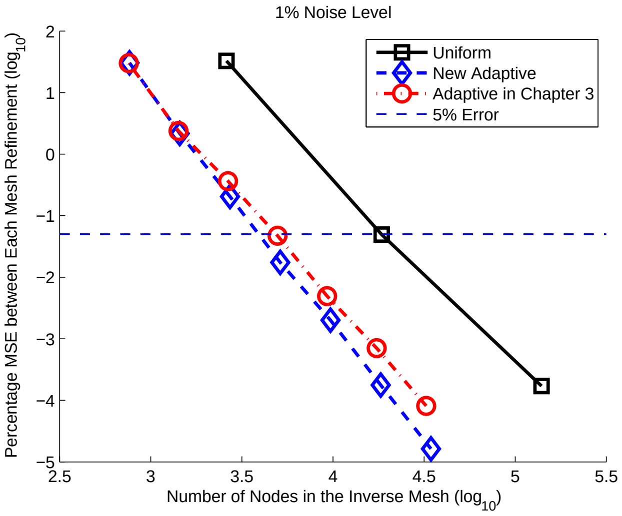


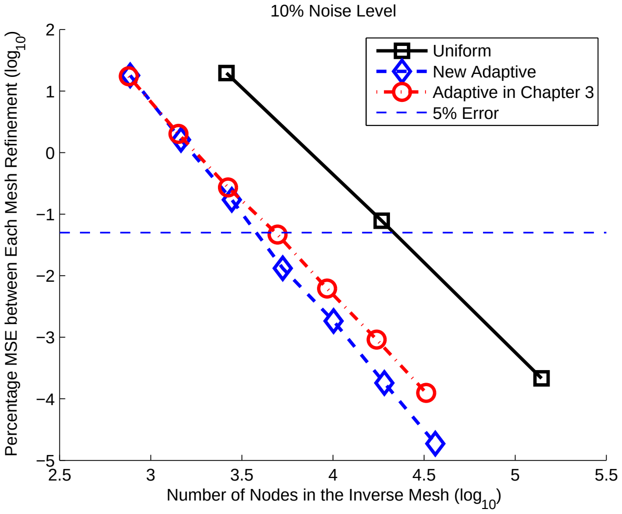
Figure 5.10: The profiles of the reconstructed images μ_{MAP}, $\mu^D_{MAP,U}$, $\mu^D_{MAP,A}$ and $\mu^D_{MAP,NA}$ along the y-axis in Simulation Study 4, Case 3, with $\mu_{axf} = 0.033$ cm^{-1} and $\kappa_f = 1 \times 10^{-4}$.

level increases. Similar to the simulation studies in Chapter 3, this error reflects the gain of the image reconstruction accuracy that can be obtained by a further mesh refinement, and we used it as an indicator to the MSE due to discretization in the reconstructed image.

Figures 5.11 and 5.12 show the change in MSE versus the number of nodes in the inverse problem mesh for different mesh generation algorithms in Simulation Study 3, Case 1 and Case 3, as well as in Simulation Study 4, Case 3. We observe that our adaptive mesh generation algorithms in this chapter have the fastest convergence rate as compared to the algorithms proposed in Chapter 3 and the uniform meshing scheme. The horizontal dash lines in Figures 5.11 and 5.12 represent the 5% error between each mesh refinement. Figures 5.11 and 5.12 show that, to achieve this accuracy level, one needs to use the uniform mesh with over 20,000 nodes and the adaptive mesh generated by the algorithm in Chapter 3 with over 5,000 nodes for the inverse problem. On the other hand, the algorithm in this chapter requires roughly 4,000 nodes in the mesh.



(a) The change in percent MSE versus the number of nodes used in the inverse problem mesh in Simulation 3, Case 1, with 1% noise.



(b) The change in percent MSE versus the number of nodes used in the inverse problem mesh in Simulation 3, Case 3, with 10% noise.

Figure 5.11: The change in percent MSE versus the number of nodes used in the inverse problem mesh in Simulation Study 3, Case 1 and Case 3. The horizontal dash line indicates the 5% error level between each mesh refinement.

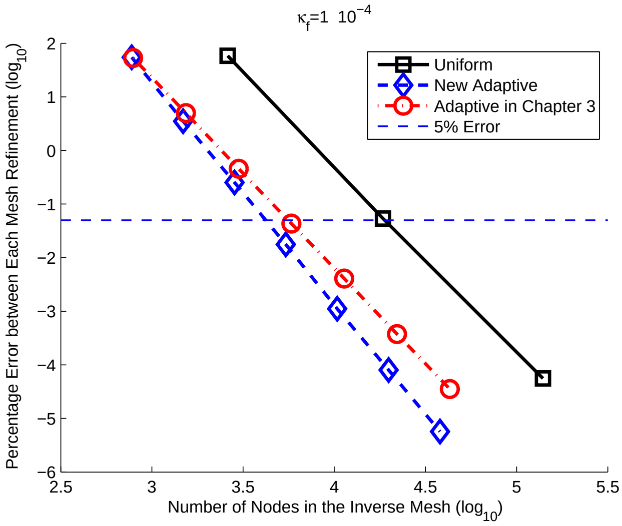


Figure 5.12: The change in percent MSE versus the number of nodes used in the inverse problem mesh in Simulation Study 4, Case 3. The horizontal dash line indicates the 5% error level between each mesh refinement.

5.6 Conclusion

Based on our error analysis in Chapter 4, we developed new adaptive mesh generation algorithms for FDOT forward and inverse problems in the presence of measurement noise. The new adaptive mesh generation algorithms take into account the noise statistics and *a priori* information on fluorophore concentration as compared with those in Chapter 3, and aim to improve the reconstruction accuracy in the mean square sense. After discussing the detailed mesh refinement processes of our new algorithms, we addressed some practical issues in their implementations, and gave several approximated error indicators which can be used to improve the computational efficiency of the mesh refinement. Finally we briefly analyzed the computational complexities of the new algorithms, and showed that they are same as those of our previous algorithms in Chapter 3 and the conventional method.

In the simulation study, we compared the adaptive mesh generation algorithms proposed in this chapter with our previous algorithms developed in Chapter 3 as well as the uniform meshing scheme. The results of the simulation study showed that our new algorithms can generate adaptive forward meshes with varying resolution not only within each mesh for a certain source-detector pair, but also across the meshes for all source-detector pairs. Furthermore, our new algorithms adaptively discretized FDOT forward and inverse problems according to the noise level as well the *a priori* information on fluorophore concentration. In the image reconstruction, we showed that the adaptive meshes generated by our previous algorithms can only provide error reduction for the bias part of the total MSE with respect to the uniform meshes. On the other hand, the meshes generated by our new algorithms can effectively reduce both bias and variance of the reconstructed images, thereby effectively reducing the total MSE as compared to other meshing schemes. We also showed that the convergence rate of the discretized inverse problem solution to the exact inverse problem solution is also improve by using our new algorithms in the FDOT reconstruction.

CHAPTER 6
PERFORMANCE EVALUATION OF ADAPTIVE FEM FOR FDOT USING EXPERIMENTAL DATA

6.1 Introduction

In Chapters 2 and 3, we analyzed the effect of discretization on the accuracy of FDOT imaging and proposed novel adaptive meshing algorithms for FDOT under the assumption that the measurements are noise-free. However, in the real FDOT reconstruction, the measurements are always corrupted by noise. Therefore, further in Chapters 4 and 5, we took the measurement noise into consideration, analyzed the effect of discretization, and proposed two new adaptive meshing algorithms for FDOT forward and inverse problems in the presence of noise. In Chapter 5, we demonstrated the implications of our error analysis and the effectiveness of our new algorithms in a numerical simulation study. In this chapter, we further apply the adaptive mesh generation algorithms proposed in Chapter 5 to the FDOT reconstruction using data obtained from a phantom experiment to show the practical advantages of our algorithms in the real FDOT reconstruction.

In the following sections, we first describe the optical phantom and FDOT imaging system setup, as well as the data collecting procedure used in the phantom experiment. To discretize and solve the FDOT forward and inverse problems, we generate the uniform meshes and the adaptive meshes by using the conventional method as well as our algorithms proposed in Chapter 5. Next, we illustrate and compare both mesh generation and image reconstruction results obtained by using our algorithms with those of the uniform and conventional adaptive meshing schemes.

6.2 Experiment Setup and Data Collection

In the phantom experiment, we used a cylindrical phantom made of silicone rubber with 2 cm diameter and 4 cm height as shown in Figure 6.1. The phantom

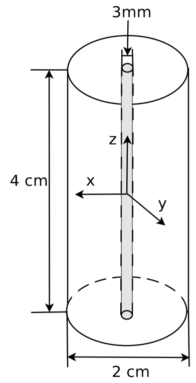


Figure 6.1: The optical phantom used in the phantom experiment.

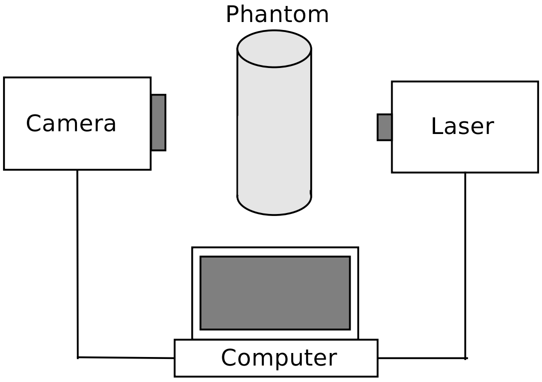


Figure 6.2: The experiment setup for the phantom experiment.

had homogeneous absorption coefficient $\mu_a = 0.2$ cm^{-1} and reduced scattering coefficient $\mu_s' = 12$ cm^{-1}. As shown in Figure 6.1, the phantom contained a hollow tube with 3 mm diameter along the central axis. The tube was filled with 1 μM Cy7 (excitation at 743 nm, emission at 767 nm) mixed with Intralipid and ink to mimic the same background optical properties of the phantom base.

The optical measurements were collected on the FDOT imaging system reported in [118], as illustrated in Figure 6.2. Specifically, a focused collimated laser beam was used as a source on one side of the phantom at 24 different positions (on a 3 × 8 grid, over a 10 mm × 15 mm field of view) sequentially. On the other side of the phantom, an electrically cooled Charged Coupled Device (CCD) camera was placed to collect the continuous wave fluorescence measurements for each source position. In the image obtained by CCD camera, we selected 36 subimages (on a 4 × 9 grid, over a 10 mm × 20 mm field of view) each with 6 × 6 pixels on the phantom as the virtual detectors, and the reading of each detector was calculated as the mean pixel value of each subimage. Figures 6.3(b) and 6.3(a) show the source and detector positions, respectively. After finishing collecting a set of measurements for all 24 source positions, the phantom is rotated by 180° to obtain another set of measurements. Therefore, the total number of measurement available for reconstruction is $24 \times 36 \times 2 = 1,728$.

(a) The detector positions illustrated in the image obtained by CCD camera.

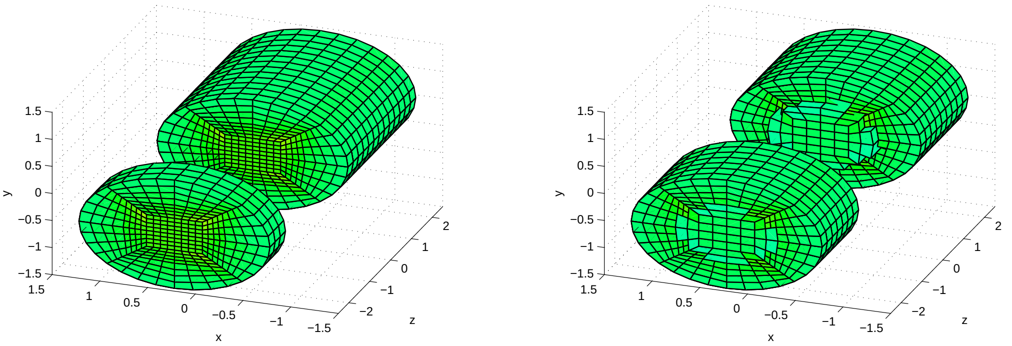
(b) The source positions illustrated in the image obtained by CCD camera.

Figure 6.3: The source and detector positions in the image obtained by CCD camera in the phantom experiment.

6.3 Mesh Generation

To numerically solve the FDOT forward and inverse problems, we used three different types of meshes with hexahedral finite elements and trilinear Lagrange basis functions to discretize the cylindrical domain Ω: uniform meshes, the adaptive meshes generated by the conventional method, and the adaptive meshes generated by our algorithms proposed in Chapter 5. We use μ_U, μ_{CA}, and μ_{NA} to denote the images reconstructed by these three types of meshes respectively. For the forward problem, the total number of nodes in all the meshes generated for all sources and detectors ranged from 600,000 to 720,000 (roughly 5,000 to 6,000 for each mesh); and for the inverse problem, it raged from 3,000 to 4,000. Figures 6.4(a) and 6.4(b) show the uniform meshes with 5,729 and 3,661 nodes generated for the forward and inverse problems, respectively.

For the adaptive meshes generated for the forward problem, Figure 6.5 shows the relationship between the number of nodes in the mesh generated by the conventional and our algorithm for a certain source (or detector), and the z-coordinate of that source (or detector) position. Note that for source or detector positions with same z-coordinate, we calculated the average number of node in the corresponding meshes. Figure 6.5 indicates that our adaptive mesh generation algorithm generated



(a) The uniform mesh with 5,729 for the forward problem.

(b) The uniform mesh with 3,661 nodes for the inverse problem.

Figure 6.4: The uniform meshes for the forward and inverse problems used in the phantom experiment. The mesh is cut through to show the mesh structure inside.

the forward meshes with varying resolution for different source-detector pairs, while the conventional method generated the forward meshes with approximately same number of nodes for all source-detector pairs. Specifically, our algorithm generated finer meshes for the source-detector pairs towards the center of the source and detector grids due to the higher SNRs of the corresponding measurements. Furthermore, our algorithm also generated finer meshes for the sources than for the detectors. Since the number of sources was less than the number of detectors in the phantom experiment, the error indicator $\varepsilon^{i}_{MAP,f}$ was expected to be higher than the error indicator $\varepsilon^{j}_{MAP,f}$ according to (5.4) and (5.5). Therefore, the adaptive meshes for the sources are generated with higher discretization level than that of the meshes for the detectors.

Additionally, both conventional and our algorithms generated the meshes for the forward and inverse problems with spatially varying resolution within each mesh. Figure 6.6 shows the adaptive meshes generated for the detector located at (-0.47,0.88,-1.12) and the source located at (0.06,-1.00,-0.30), and Figure 6.7 shows the adaptive meshes generated for the inverse problem. Clearly, the meshes shown in Figures 6.6(c), 6.6(d), and 6.7(b) were generated by the conventional method only based on the corresponding forward or inverse problem solution. On the other hand, as shown in Figures 6.6(a), 6.6(b), and 6.7(a), our algorithms took the in-

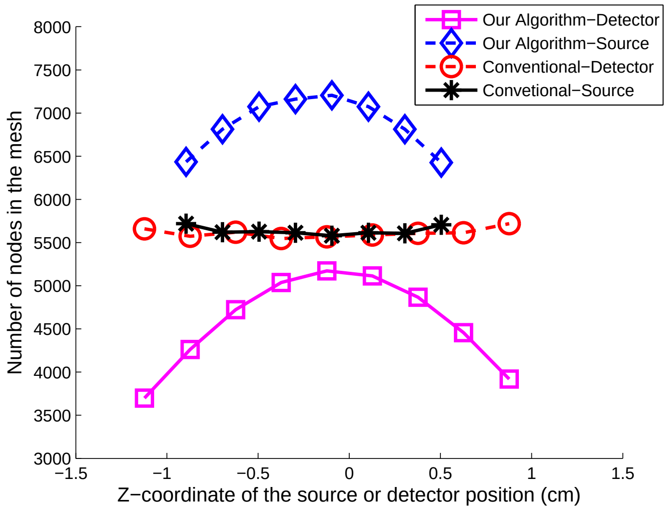


Figure 6.5: The relationship between the number of nodes in the forward adaptive mesh for a certain source (or detector) and the z-coordinate of the source (or detector) position.

terdependence of forward and inverse problems into account, and refined the mesh based on the solutions of both problems.

6.4 Experiment Results

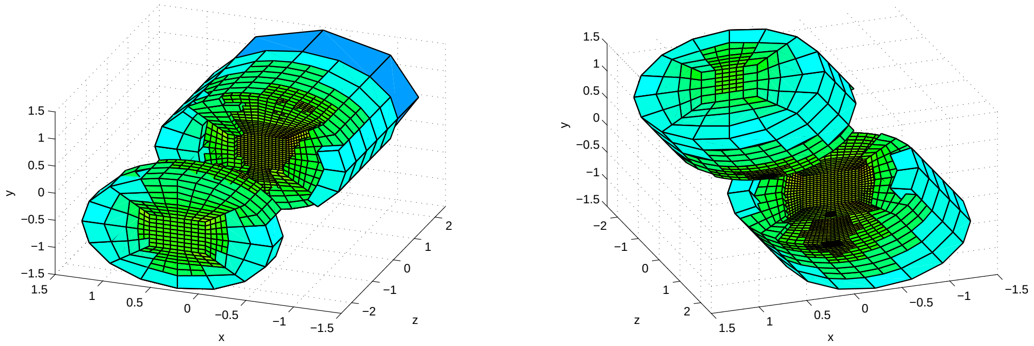
To reconstruct the fluorophore concentration image, we formulated the FDOT inverse problem as the MAP estimator of the fluorophore concentration given in Chapter 4. For the noise model, we considered the same Gaussian approximation to the short-noise model:

$$\sigma_k^2 = \alpha|\Gamma_k|,$$

where σ_k^2 is the variance of the noise in the k^{th} measurement Γ_k, $k = 1, \ldots, 1,728$, and α is a constant. We estimated α by calculating the pixel variance in a number of homogeneous regions in the images obtained by the CCD camera.

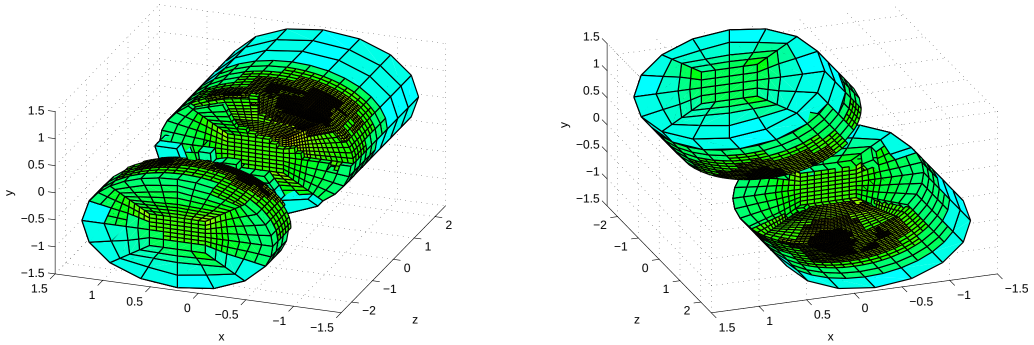
We considered an *a priori* model for the fluorophore concentration $\mu(\boldsymbol{r})$ given by

$$\kappa(\boldsymbol{r}) = \begin{cases} \kappa_f & \boldsymbol{r} \in \Omega_f \\ \kappa_b & \boldsymbol{r} \in \Omega_b \end{cases},$$



(a) The adaptive mesh with 3,745 nodes generated by our algorithm proposed in Chapter 5 for the detector located at (-0.47,0.88,-1.12).

(b) The adaptive mesh with 7,128 nodes generated by our new algorithm proposed in Chapter 5 for the source located at (0.06,-1.00,-0.30).



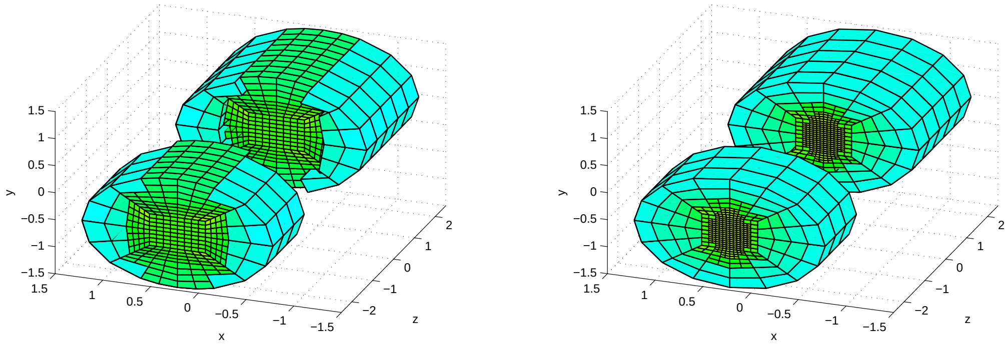
(c) The adaptive mesh with 5,794 nodes generated by the conventional method for the detector located at (-0.47,0.88,-1.12).

(d) The adaptive mesh with 5,831 nodes generated by the conventional method for the source located at (0.06,-1.00,-0.30).

Figure 6.6: Examples of the adaptive meshes for the forward problem used in the phantom experiment. The mesh is cut through to show the mesh structure inside.

where κ_f and κ_b are constant, $\Omega_f \cup \Omega_b = \Omega$, $\Omega_f \cap \Omega_b = \emptyset$ and Ω_f is the region of the hollow tube that contains the fluorophore inside the phantom. We chose $\kappa_f = 2 \times 10^{-5}$ and $\kappa_b = 1 \times 10^{-5}$ according to the higher variance of the fluorophore concentration in the foreground Ω_f than that in the background Ω_b.

To obtain the exact solution of the forward and inverse problems, we discretized the domain using a fine mesh with 91,377 nodes, and used the resulting image, denoted by μ, as the baseline for calculating the error due to discretization. After reconstructing all the images, we calculated the $L^2(\Omega)$ norms of the differences between the baseline image and the images reconstructed by coarse meshes to



(a) The adaptive mesh with 3,784 nodes generated by our algorithm proposed in Chapter 5 for the inverse problem.

(b) The adaptive mesh with 3,969 nodes generated by the conventional method for the inverse problem.

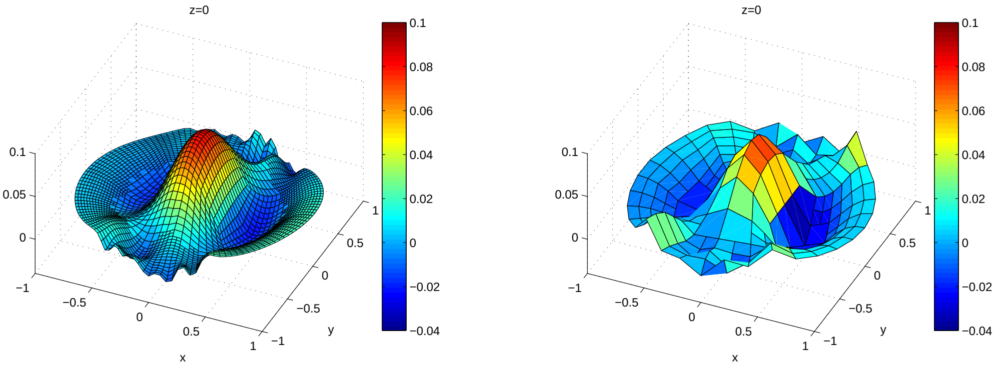
Figure 6.7: The adaptive meshes for the inverse problem used in the phantom experiment. The mesh is cut through to show the mesh structure inside.

Table 6.1: $L^2(\Omega)$ norm of the error due to discretization in the images reconstructed by using different meshes in the phantom experiment.

	Absolute Value	Percentage Value
$\|\mu - \mu_U\|_0$	6.096×10^{-3}	100%
$\|\mu - \mu_{CA}\|_0$	5.134×10^{-3}	84.23%
$\|\mu - \mu_{NA}\|_0$	3.283×10^{-3}	53.87%

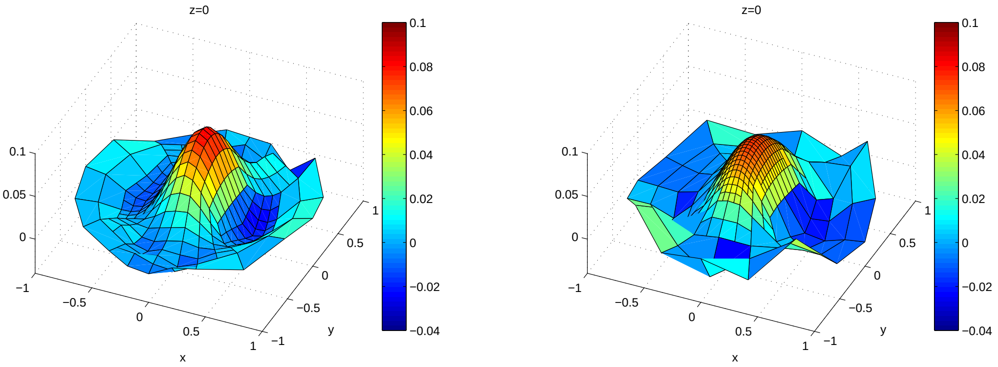
evaluate the reconstruction accuracy. We tabulated the results in Table 6.1. The results show that, as compared to the image reconstructed using uniform meshes, the error due to discretization is reduced by approximately 45% when our algorithms are used in FDOT reconstruction. On the other hand, the conventional method can only provide approximately 15% error reduction with respect to the uniform meshing scheme.

Figure 6.8 shows the cross-section of the fluorophore concentration images at $z = 0$ plane reconstructed by using different types of meshes. The baseline image μ is shown in Figure 6.8(a). As shown in Figures 6.8(b), 6.8(c) and 6.8(d), the shape of the fluorophore heterogeneity is better resolved in μ_{NA} and μ_{CA} with spatially varying resolution than in μ_U. We also observe that μ_{NA} is the closest one to μ in all images reconstructed using coarse meshes. More specifically, μ_A has higher response in the center of the fluorophore heterogeneity and lower background variation as



(a) The baseline image.

(b) The image reconstructed using the uniform meshes.



(c) The image reconstructed using the adaptive meshes generated by our algorithms proposed in Chapter 5.

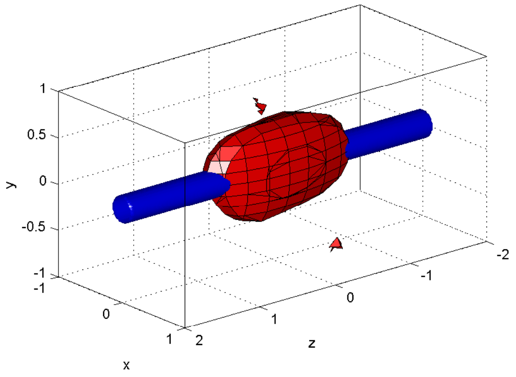
(d) The image reconstructed using the adaptive meshes generated by the conventional method.

Figure 6.8: The cross-sectional reconstruction result of phantom experiment on plane $z = 0$.

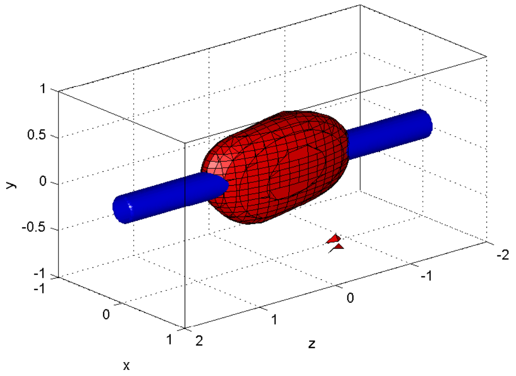
compared to those of μ_U and μ_{CA}.

Figure 6.9 shows the reconstructed surface of the fluorophore heterogeneity at the half its maximum concentration value obtained by using different meshes in image reconstruction. We observe some small objects reconstructed away from the true heterogeneity (indicated by the blue tubes in Figure 6.9) in Figures 6.9(a) and 6.9(b), which represent the background variation with value higher than the half maximum fluorophore concentration value in the reconstructed image. On the other hand, the result obtained by using our algorithms has a clean background, and a more smooth reconstructed surface, as well as a more compact shape close to the true tube-shaped fluorophore heterogeneity as shown in Figure 6.9(c).

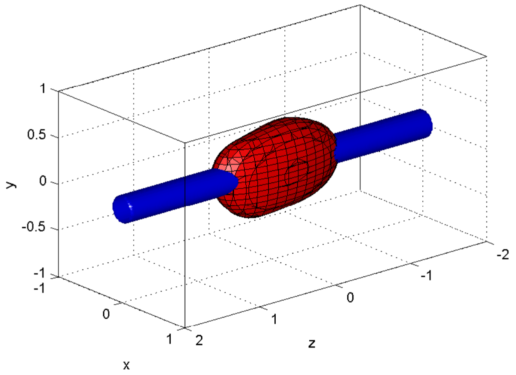
The results in Table 6.1 and Figures 6.8 and 6.9 demonstrate the improvements



(a) The 3D fluorophore heterogeneity reconstructed using uniform meshes.



(b) The 3D fluorophore heterogeneity reconstructed using the adaptive meshes generated by our algorithms proposed in Chapter 5.



(c) The 3D fluorophore heterogeneity using the adaptive meshes generated by the conventional method.

Figure 6.9: The 3D fluorophore heterogeneity reconstructed in the phantom experiment. The surface is reconstructed at the half maximum value of the reconstructed fluorophore concentration. The blue tube indicates the true fluorophore heterogeneity.

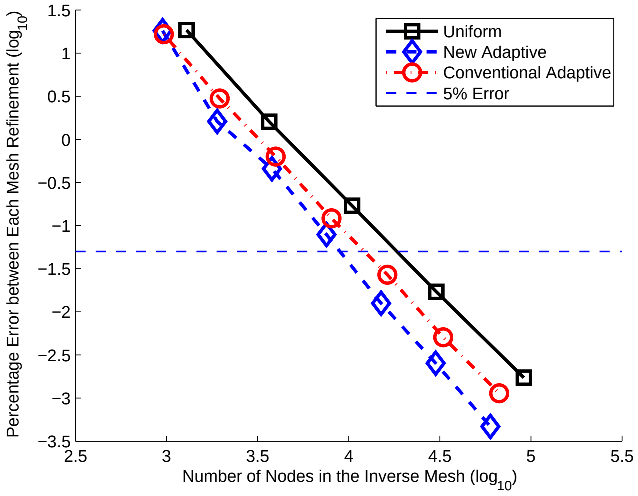


Figure 6.10: The change in percent $L^2(\Omega)$ norm of error versus the number of nodes used in the inverse problem mesh in the phantom experiment. The horizontal dash line indicates the 5% error level between each mesh refinement.

in reconstruction accuracy, and small target detectability, as well as image resolution of the reconstructed optical images, when our adaptive meshing algorithms are used in FDOT reconstruction.

We finally studied the convergence of the discretized inverse problem solution to the exact solution when different types of meshes were used in the phantom experiment. We used the similar method as we used in the simulation studies in Chapters 3 and 5. Figure 6.10 shows the change in the $L^2(\Omega)$ norm of error versus the number of nodes used in the inverse problem mesh for different adaptive mesh generation algorithms in the phantom experiment. We observe that our adaptive mesh generation algorithms again have the fastest convergence rate among all meshing schemes. The horizontal dash line in Figure 6.10 represents the 5% error between each mesh refinement. The result shows that, to achieve this accuracy level, one only needs to use our adaptive mesh with roughly 8,000 nodes for the inverse problem, as compared to the conventional adaptive mesh with over 12,000 nodes and the uniform mesh with over 19,000 nodes.

6.5 Conclusion

In this chapter, we demonstrated the practical advantages of our adaptive mesh generation algorithms proposed in Chapter 5 in the real FDOT reconstruction using the data collected from a phantom experiment. In the real FDOT reconstruction case, the measurement noise is an important factor that affects the reconstruction accuracy together with the problem discretization. Taking into account the interdependence of the FDOT forward and inverse problems as well as the measurement noise, our adaptive mesh generation algorithms can adaptively refine the meshes according to the factors, such as the source detector configuration, their positions with respect to the fluorophore heterogeneity, the noise statistics and *a priori* information on the fluorophore concentration. The reconstruction results showed that our new algorithms can effectively improve the accuracy, small target detectability, and resolution, as well as the convergence rate of the reconstructed optical images using real data.

CHAPTER 7
CONCLUSION AND FUTURE WORK

7.1 Conclusion

In this thesis, we analyzed the effect of discretization on the accuracy of FDOT imaging when the measurements are noise-free as well as in the presence of measurement noise. We presented our error analysis results in the form of several upper bounds for the error due to discretization in the reconstructed optical images. These error bounds take into account the mutual dependence of the forward and inverse problems, and show the factors that determine the extent to which the discretizations can affect the accuracy of FDOT reconstruction. These factors include the number of sources and detectors, their configuration and their positions with respect to the fluorophore heterogeneity, and the formulation of the inverse problem when the measurements are noise-free. They also include the noise statistics and *a priori* information on fluorophore concentration when measurements are contaminated with noise. The two most important implications of our error analysis can be summarized as follows:

- One has to take into account the mutual dependence of the forward and inverse problems, and discretize either forward or inverse problem of FDOT based on the solutions of both problems.

- In the presence of measurement noise, one has to additionally take into account the noise statistics and *a priori* information on fluorophore concentration, while discretizing either the forward or inverse problem of FDOT.

The error analysis in Chapters 2 and 4 motivates the development of adaptive mesh generation algorithms for FDOT. In Chapters 3 and 5, we presented four novel adaptive mesh generation algorithms for the forward and inverse problems of FDOT for the noise-free case and for the case when measurements are contaminated by noise, respectively. All of these new adaptive mesh generation algorithms generate meshes with spatially varying resolutions based on the solutions of both problems. Furthermore, the two algorithms proposed in Chapter 4 can additionally refine the meshes according to the noise statistics and *a priori* information on the fluorophore concentration.

The numerical simulation studies in Chapters 3 and 5 validated our error analysis and demonstrated the advantages of our new adaptive mesh generation algorithms in FDOT reconstruction. In these simulation studies, we showed that the adaptive mesh generation algorithms developed in this thesis can effectively improve the accuracy, resolution, and small target detectability of FDOT reconstruction, as well as the convergence rate of the discretized solution to the exact inverse problem solution, as compared to the uniform and the conventional adaptive meshing schemes. In Chapter 6, we also showed the practical use of our adaptive mesh generation algorithms in FDOT reconstruction using data from a phantom experiment. The reconstruction results using the phantom data were consistent with the results of our simulation studies, and demonstrated the effectiveness of our algorithms in real data.

In Chapter 4, we formulated our FDOT inverse problem based on the independent Gaussian models for both measurement noise and fluorophore concentration. In some situations, these models may not result in the optimal FDOT reconstruction due to the mismatch between the models and the underlying natures of the measurement noise and fluorophore concentration. However, the error analysis and adaptive mesh generation algorithms presented in the thesis can be extended to the case when different noise and *a priori* fluorophore concentration models are considered. Particularly, we discussed the case when more general Gaussian models are used for the measurement noise as well as the fluorophore concentration, and presented the corresponding error bounds for this case.

Finally, we summarize the contributions of this thesis:

- The discretization error analysis in this thesis takes into account the interdependence of FDOT forward and inverse problems, as well as the impact of measurement noise. The resulting error bounds explicitly show the effect of problem discretizations on FDOT imaging and identify the key factors that determine the extent to which the discretization can affect the FDOT reconstruction accuracy. We note that this part of the work is the first study in the literature that takes into account measurement noise and provides a rigorous analysis on the error due to discretization in FDOT imaging.

- The adaptive mesh generation algorithms developed based on our error analysis can effectively reduce the error due to discretization in the reconstructed optical images as compared to the uniform and conventional adaptive meshing schemes. We note that the simulation and phantom studies in this thesis are the first studies in the literature that provide comprehensive performance evaluation and comparison among different adaptive meshing methods in FDOT imaging.
- The error analysis and adaptive mesh generation algorithms in this thesis are not limited to FDOT imaging problem, and can be adapted to other inverse parameter estimation problems such as bioluminescence tomography, electrical impedance tomography, and microwave imaging.

7.2 Future Work

Based on the perspectives, methods, and algorithms developed in this thesis, we propose some future work in the following directions:

- In this thesis, we focus on estimating the fluorophore concentration using the CW method. However, the approach we presented in this thesis is not limited to the single frequency FDOT imaging system, and can be extended to FD and TD methods to estimate both concentration and lifetime of the fluorophore.
- In FDOT reconstruction, the regularization and *a priori* model of fluorophore concentration cause the bias in the reconstructed image, which is another type of error in addition to the errors due to discretization and measurement noise. The regularization parameters and variance in the *a priori* fluorophore concentration model provide a way to balance this bias and the error due to measurement noise. Our error bounds given in Chapters 2 and 4 are the functions

of these parameters. In the development of adaptive mesh generation algorithms, we assume that the appropriate values for regularization parameters and variance of fluorophore concentration are determined *a priori*. However, since our error bounds involve these parameters, it is possible to study the interplay between the choice of these parameters and the mesh size, and adaptively select them in an attempt to reduce the overall error in reconstructed optical images.

- In the presence of measurement noise, we formulate the FDOT imaging problem as an estimation problem to estimate the fluorophore concentration using the noise-contaminated data. We access the quality of FDOT reconstruction by the MSE due to discretization in the fluorophore concentration estimate. Alternatively, the FDOT imaging problem can also be formulated as a detection problem where the objective is to detect the fluorophore heterogeneity from the background variation in the reconstructed image. Therefore, a detection figure of merit, such as SNR of the test-statistic, can be chosen to evaluate the quality of the FDOT reconstruction. A study on the effect of discretization on the SNR of the reconstructed optical image can motivate the development of the adaptive mesh generation algorithm that can be used to maximize the detectability of the FDOT imaging system with the constraint of the computation complexity.

BIBLIOGRAPHY

[1] D. A. Boas, D. H. Brooks, E. L. Miller, C. A. Dimarzio, M. Kilmer, R. J. Gaudette, and Q. Zhang. Imaging the body with diffuse optical tomography. *IEEE Signal Proc. Magazine*, 18(6):57–75, 2001.

[2] A. P. Gibson, J. C. Hebden, and S. R. Arridge. Recent advances in diffuse optical imaging. *Phys. Med. Biol.*, 50:R1–43, 2005.

[3] B. Chance. Optical method. *Annu. Rev. Biophys. Biophys. Chem.*, 20:1–28, 1991.

[4] T. O. McBride, B. W. Pogue, E. D. Gerety, S. B. Poplack, U. L. Öterberg, and K. D. Paulsen. Spectroscopic diffuse optical tomography for the quantitative assessment of hemoglobin concentration and oxygen saturation in breast tissue. *Appl. Opt.*, 38(25):5480–5490, 1999.

[5] T. McBride, B. W. Pogue, S. Jiang, U. L. Osterberg, and K. D. Paulsen. A parallel-detection frequency-domain near-infrared tomography system for hemoglobin imaging of the breast in vivo. *Rev. Sci. Instrum.*, 72:1817–24, 2001.

[6] D. Grosenick, H. Wabnitz, K. T. Moesta, J. Mucke, M. Möller, C. Stroszczynski, J. Stössel, B. Wassermann, P. M. Schlag, and H. Rinneberg. Concentration and oxygen saturation of haemoglobin of 50 breast tumours determined by time-domain optical mammography. *Phys. Med. Biol.*, 49:1165–81, 2004.

[7] B. W. Pogue, S. Jiang, H. Dehghani, C. Kogel, S. Soho, S. Srinivasan, X. Song, T. D. Tosteson, S. P. Poplack, and K. D. Paulsen. Characterization of hemoglobin, water and nir scattering in breast tissue: analysis of intersubject variability and menstrual cycle changes. *J. Biomed. Opt.*, 9:541–52, 2003.

[8] T. D. Yates, J. C. Hebden, A. P. Gibson, N. L. Everdell, D. T. Delpy, S. R. Arridge, M. Douek, and W. Chicken. Clinical results from a 32-channel time resolved system used to image the breast. page WF18. Miami, 2004.

[9] V. Ntziachristos, A. G. Yodh, M. Schnall, and B. Chance. Concurrent mri and diffuse optical tomography of breast after indocyanine green enhancement. pages 2767–72. USA, 2000.

[10] X. Intes, J. Ripoll, Y. Chen, S. Nioka, A. G. Yodh, and B. Chance. In vivo continuous-wave optical breast imaging enhanced with indocyanine green. *Med. Phys.*, 30:1039–47, 2003.

[11] G. Strangman, D. A. Boas, and J. P. Sutton. Non-invasive neuroimaging using near-infrared light. *Biological Psychiatry*, 52:679–693, 2002.

[12] A. Villringer and B. Chance. Non-invasive optical spectroscopy and imaging of human brain function. *Trends in Neurosciences*, 20(10):435–442, 1997.

[13] B. Chance, S. Nioka, and Y. Chen. Shining new light on brain function. *OE Mag.*, 3:16–19, 2003.

[14] J. C. Hebden. Advances in optical imaging of the newborn infant brain. *Psychophysiology*, 40:501–10, 2003.

[15] R. Weissleder and V. Ntziachristos. Shedding light onto live melecular targets. *Nature Medicine*, 9:123–128, 2003.

[16] V. Ntziachristos, J. Ripoll, L. V. Wang, and R. Weissleder. Looking and listening to light: the evolution of whole-body photonic imaging. *Nature Biotechnology*, 23:313–320, 2005.

[17] C. Bremer, V. Ntziachristos, and R. Weissleder. Optical-based molecular imaging: contrast agents and potential medical applications. *Eur. Radiol.*, 13:231–43, 2003.

[18] W. C.W. Chan, D. J. Maxwell, X. Gao, R. E. Bailey, M. Han, and S. Nie. Luminescent quantum dots for multiplexed biological detection and imaging. *Opin. Biotechnol.*, 13:40–6, 2002.

[19] S. R. Cherry. In vivo molecular and genomic imaging: new challenges for imaging physics. *Phys. Med. Biol.*, 49:R13–48, 2004.

[20] J. R. Lakowicz. *Principles of fluorescence spectroscopy.* Plenum, New York, 1983.

[21] V. Ntziachristos and R. Weissleder. Charge-coupled-device based scanner for tomography of fluorescent nearinfrared probes in turbid media. *Med. Phys.*, 29:803–9, 2002.

[22] E. E. Graves, J. Ripoll, R. Weissleder, and V. Ntziachristos. A submillimeter resolution fluorescence molecular imaging system for small animal imaging. *Med. Phys.*, 30:901–11, 2003.

[23] G. Zacharakis, J. Ripoll, R. Weissleder, and V. Ntziachristos. Fluorescent protein tomography scanner for small animal imaging. *IEEE Tran. on Mecial Imaging*, 24(7):878–85, 2005.

[24] R. B. Schulz, J. Ripoll, and V. Ntziachristos. Experimental fluorescence tomography of tissues with noncontact measurements. *IEEE Transactions on Medical Imaging*, 23(4):492–500, 2004.

[25] H. Jiang. Frequency-domain fluorescent diffusion tomography: a finite-element-based algorithm and simulations. *Appl. Opt.*, 37(22):5337–43, 1998.

[26] R. Roy and E. M. Sevick-Muraca. Three-dimensional unconstrained and constrained image-reconstruction techniques applied to fluorescence, frequency-domain photon migration. *Appl. Opt.*, 40(13):2206–15, 2001.

[27] A. B. Milstein, J. J. Stott, S. Oh, D. A. Boas, R. P. Millane, C. A. Bouman, and K. J. Webb. Fluorescence optical diffusion tomography using multiple-frequency data. *J. Opt. Soc. Am. A*, 21(6):1035–49, 2004.

[28] S. Lam, F. Lesage, and X. Intes. Time domain fluorescent diffuse optical tomography: analytical expressions. *Opt. Express*, 13(7):2263–75, 2005.

[29] S. B. Raymond, G. Boverman, D. A. Boas, and B. J. Bacskai. Time resolved fluorescence tomography of turbid media based on lifetime contrast. *Opt. Express*, 14(25):12255–70, 2006.

[30] A. T.N. Kumar, S. B. Raymond, A. K. Dunn, B. J. Bacskai, and D. A. Boas. A time domain fluorescence tomography system for small animal imaging. *IEEE Transactions on Medical Imaging*, 27(8):1152–63, 2008.

[31] K. Chen, L. T. Perelman, Q. Zhang, R. R. Dasari, and M. S. Feld. Optical computed tomography in a turbid medium using early arriving photons. *J. Biomed. Opt.*, 5:144–54, 2000.

[32] R. B. Schulz, J. Ripoll, and V. Ntziachristos. Experimental fluorescence tomography of arbitrarily shaped diffuse objects using non-contact measurements. *Opt. Lett.*, 28:1701–03, 2003.

[33] J. Ripoll, R. B. Schulz, and V. Ntziachristos. Free-space propagation of diffuse light: theory and experiment. *Phys. Rev. Lett.*, 91:103901–04, 2003.

[34] A. Ishimaru. *Wave propagation and scattering in random media.* IEEE Press, New York, 1997.

[35] O. Dorn. A transport-backtransport method for optical tomography. *Inverse Problems*, 14(2):1107–1130, 1998.

[36] S. R. Arridge. Optical tomography in medical imaging. *Inverse Problems*, 15:R41–93, 1999.

[37] M. S. Patterson and B. W. Pogue. Mathematical model for time-resolved and frequency-domain fluorescence spectroscopy in biological tissues. *Appl. Opt.*, 33:1963–74, 1994.

[38] E. M. Sevick-Muraca, G. Lopez, J. S. Reynolds, T. L. Troy, and C. L. Hutchinson. Fluorescence and absorption contrast mechanisms for biomedical optical imaging using frequency-domain techniques. *Photochem. Photobiol.*, 66:56–64, 1997.

[39] H. Dehghani, D.T. Delpy, and S.R. Arridge. Photon migration in nonscattering tissue and the effects on image reconstruction. *Phys. Med. Biol.*, 44:2297–2906, 1999.

[40] J. Riley, H. Dehghanni, M. Schweiger, S. Arridge, J. Ripoll, and M. Nieto-Vesperinas. Optical tomography in the presence of void regions. *Opt. Express*, 7(13).

[41] D. A. Boas, M. A. O'Leary, B. Chance, and A. G. Yodh. Scattering of diffuse photon density waves by spherical inhomogeneities within turbid media: analytic solution and application. *Proc. Natl. Acad. Sci. USA*, 91:4887–91, 1994.

[42] F. Martelli, A. Sassaroli, Y. Yamada, and G. Zaccanti. Analytical approximate solutions of the time-domain diffusion equation in layered slabs. *J. Opt. Soc. Am. A*, 19:71–80, 2002.

[43] J. Ripoll, V. Ntziachristos, R. Carminati, and M. Nieto-Vesperinas. Kirchhoff approximation for diffusive waves. *Phys. Rev. E*, 64:051917, 2001.

[44] D. A. Boas, J. P. Culver, J. J. Stott, and A. K. Dunn. Three dimensional monte carlo code for photon migration through complex heterogeneous media including the adult human head. *Opt. Express*, 10:159–69, 2002.

[45] E. Okada and D. T. Delpy. Near-infrared light propagation in an adult head model: I. modeling of low-level scattering in the cerebrospinal fluid layer. *Appl. Opt.*, 42:2906–14, 2003.

[46] T. Hayashi, Y. Kashio, E. Okada, and A. K. Dunn. Hybrid monte carlo-diffusion method for light propagation in tissue with a low-scattering region. *Appl. Opt.*, 42:2888–96, 2003.

[47] V. Chernomordik, D. W. Hattery, D. Grosenick, H. Wabnitz, H. Rinneberg, K. T. Moesta, P. M. Schlag, and A. H. Gandjbakhche. Quantification of optical properties of a breast tumour using random walk theory. *J. Biomed. Opt.*, 7:80–7, 2002.

[48] H. Dehghani, B. Brooksby, K. Vishwanath, B. W. Pogue, and K. D. Paulsen. The effect of internal refractive index variation on near-infrared optical tomography: a finite elementmodelling approach. *Phys. Med. Biol.*, 48:2713–27, 2003.

[49] B. Brooksby, H. Dehghani, B. W. Pogue, and K. D. Paulsen. Near infrared (nir) tomography breast image reconstruction with a priori structural information from mri: algorithm development for reconstructing heterogeneities. *IEEE Journal of Selected Topics in Quantum Electronics*, 9:199–209, 2003.

[50] A. P. Gibson, J. Riley, M. Schweiger, J. C. Hebden, S. R. Arridge, and D. T. Delpy. A method for generating patient-specific finite element meshes for head modelling. *Phys. Med. Biol.*, 48:481–95, 2003.

[51] H. Dehghani, M. M. Doyley, B. W. Pogue, S. Jiang, J. Geng, and K. D. Paulsen. Breast deformation modelling for image reconstruction in near infrared optical tomography. *Phys. Med. Biol.*, 49:1131–45, 2004.

[52] M. A. Oleary, D. A. Boas, X. D. Li, B. Chance, and A. G. Yodh. Fluorescence lifetime imaging in turbid media. *Opt. Lett.*, 21:158–60, 1996.

[53] J. W. Chang, H. L. Graber, and R. L. Barbour. Luminescence optical tomography of dense scattering media. *J. Opt. Soc. Am. A*, 14:288–99, 1997.

[54] R. J. Gaudette, D. H. Brooks, C. A. DiMarzio, M. 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. *Phys. Med. Biol.*, 45:1051–70, 2000.

[55] J. Kaipio and E. Somersalo. *Computational and statistical inverse problems*, volume 160 of *Applied Mathematical Sciences*. Springer-Verlag, New York, 2005.

[56] J. P. Kaipio, V. Kolehmainen, M. Vauhkonen, and E. Somersalo. Inverse problems with structural prior information. *Inverse Problems*, 15:713–29, 1999.

[57] A. B. Milstein, S. Oh, K. J. Webb, C. A. Bouman, Q. Zhang, D. A. Boas, and R. P. Millane. Fluorescence optical diffusion tomography. *Appl. Opt.*, 42(16):3081–94, 2003.

[58] A. Joshi, W. Bangerth, and E. M. Sevick-Muraca. Adaptive finite element based tomography for fluorescence optical imaging in tissue. *Opt. Express*, 12(22):5402–5417, 2004.

[59] W. Bangerth and A. Joshi. Adaptive finite element methods for the solution of inverse problems in optical tomography. *Inverse Problems*, 24:034011, 2008.

[60] 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.

[61] 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.

[62] M. Ainsworth and J. T. Oden. A unified approach to a posteriori error estimation using elemental residual methods. *Numerische Mathematik*, 65:23–50, 1993.

[63] I. Babuška and W. C. Rheinboldt. Error estimates for adaptive finite element computations. *SIAM Journal on Numerical Analysis*, 15:736–754, 1978.

[64] 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.

[65] R. E. Bank and A. Weiser. Some a posterior error estimators for elliptic partial differential equations. *Mathematics of Computation*, 44:283–301, 1985.

[66] 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.

[67] R. Verfurth. *A review of a posteriori error estimation and adaptive mesh refinement techniques.* Teubner-Wiley, 1996.

[68] W. Bangerth. A framework for the adaptive finite element solution of large-scale inverse problems. *SIAM J. Sci. Comput.*, 30(6):2965–89, 2008.

[69] W. Bangerth. *Adaptive finite element methods for the identification of distributed parameters in partial differential equations.* PhD thesis, University of Heidelberg, 2002.

[70] R. Li, W. Liu, H. Ma, and T. Tang. Adaptive finite element approximation for distributed elliptic optimal control problems. *SIAM J. Control Optim.*, 41:1321–49, 2002.

[71] R. Becker and B. Vexler. A posteriori error estimation for finite element discretization of parameter identification problems. *Numer. Math.*, 96:435–59, 2003.

[72] D. Meidner and B. Vexler. Adaptive space-time finite element methods for parabolic optimization problems. *SIAM J. Control Optim.*, 46:116–42, 2007.

[73] B. Vexler and W. Wollner. Adaptive finite elements for elliptic optimization problems with control constraints. *SIAM J. Control Optim.*, 47:509–34, 2008.

[74] L. Beilina and C. Johnson. A posteriori error estimation in computational inverse scattering. *Mathematical Models and Methods in Applied Sciences*, 15:23–37, 2005.

[75] L. Beilina and C. Johnson. Adaptive finite element/difference method for inverse elastic scattering waves. *Applied and Computational Mathematics*, 2:158–174, 2003.

[76] X. Gu, Y. Xu, and H. Jiang. Mesh-based enhancement schemes in diffuse optical tomography. *Med. Phys.*, 30(5):861–869, 2003.

[77] M. Huang and Q. Zhu. Dual-mesh optical tomography reconstruction method with a depth correction that uses a priori ultrasound information. *Appl. Opt.*, 43(8):1654–1662, 2004.

[78] 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.

[79] 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.

[80] A. Joshi, W. Bangerth, K. Hwang, J. C. Rasmussen, and E. M. Sevick-Muraca. Fully adaptive fem based fluorescence optical tomography from time-dependent measurements with area illumination and detection. *Med. Phys.*, 33(5):1299–1310, 2006.

[81] J. Lee, A. Joshi, and E. Sevick-Muraca. Fully adaptive finite element based tomography using tetrahedral dual-meshing for fluorescence enhanced optical imaging in tissue. *Opt. Express*, 15(11):6955–6975, 2007.

[82] J. Lee, A. Joshi, and E. Sevick-Muraca. Fast intersections on nested tetrahedrons (fint): An algorithm for adaptive finite element based distributed parameter estimation. *Journal of Computational Physics*, 227:5778–5798, 2008.

[83] V. Ntziachristos and R. Weissleder. Experimental three-dimensional fluorescence reconstruction of diffuse media by use of a normalized born approximation. *Opt. Lett.*, 26(12):893–895, 2001.

[84] F. Fedele, J.P. Laible, and M.J. Eppstein. Coupled complex adjoint sensitivities for frequency-domain fluorescence tomography: theory and vectorized implementation. *J. Comp. Phys.*, 187:597–619, 2003.

[85] M. J. Eppstein, F. Fedele, J.P. Laible, C. Zhang, A. Godavarty, and E.M. Sevick-Muraca. A comparison of exact and approximate adjoint sensitivies in fluorescence tomography. *IEEE Transactions on Medical Imaging*, 22(10):1215–1223, 2003.

[86] R. Kress. *Linear integral equations*, volume 82 of *Applied Mathematical Sciences*. Springer-Verlag, second edition, 1999.

[87] M. Guven, B. Yazici, and V. Ntziachristos. Optical fluorescence tomography with a priori information. In *Proc. of SPIE*, volume 6431, 2007.

[88] N. P. Galatsanos and A. K. Katsaggelos. Methods for choosing the regularization parameter and estimatingthe noise variance in image restoration and their relation. *IEEE Transactions on Image Processing*, 1:322–336, 1992.

[89] M. Hanke and T. Raus. A general heuristic for choosing the regularization parameter in ill-posed problems. *SIAM Journal on Scientific Computing*, 17:956–972, 1996.

[90] R. Molina, A. K. Katsaggelos, and J. Mateos. Bayesian and regularization methods for hyperparameter estimation in image restoration. *IEEE Transactions on Image Processing*, 8:231–246, 1999.

[91] M. Heath, G. Golub, and G. Wahba. Generalized cross-validation as a method for choosing a good ridge paramete. *Technometrics*, 21:215C223, 1979.

[92] P. Hansen and D. O'Learly. The use of the l-curve in the regularization of discrete ill-posed problems. *SIAM J. Comput.*, 14:1487C1503, 1993.

[93] G. Aubert and P. Kornprobst. *Mathematical Problems in Image Processing.* Springer Verlag, 2002.

[94] S. C. Brenner and L. R. Scott. *The Mathematical Theory of Finite Element Methods.* Springer Verlag, 2002.

[95] D. Daners. Robin boundary value problems on arbitrary domains. *Transactions of the American Mathematical Society*, 352(9):4207–4236, 2000.

[96] O. Axelsson and V. Barker. *Finite Element Solution of Boundary Value Problems, Theory and Computation.* Academic Press, Orlando, FL, 1984.

[97] M. C. Rivara. Mesh refinement processes based on the generalized bisection of simplices. *SIAM Journal on Numerical Analysis*, 21(3):604–613, 1984.

[98] W. Bangerth, R. Hartmann, and G. Kanschat. deal.II — a general-purpose object-oriented finite element library. *ACM Trans. Math. Softw.*, 33(4), 2007.

[99] I. M. Smith and D. V. Griffiths. *Programming the Finite Element Method.* John Wiley & Sons, 4 edition, 2004.

[100] J. C. Ye, K. J. Webb, and C. A. Bouman. Optical diffusion tomography by iterative-coordinate-descent optimization in a bayesian framework. *J. Opt. Soc. Am. A*, 16(10):2400–12, 1999.

[101] V. H. Poor. *An introduction to signal detection and estimation.* Springer, 2 edition, 1994.

[102] R. H. Kingston. *Detection of optical and infrared radiation.* Spring-Verlag, New York, 1978.

[103] B. Brooksby, S. Jiang, H. Dehghani, B. W. Pogue, and K. D. Paulsen. Combining near-infrared tomography and magnetic resonance imaging to study in vivo breast tissue: implementation of a laplacian-type regularization to incorporate magnetic resonance structure. *J. Biomed. Opt.*, 10(5):051504, 2005.

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

[105] X. Intes, C. Maloux, M. Guven, B. Yazıcı, and B. Chance. Diffuse optical tomography with physiological and spatial a priori constraints. *Phys. Med. Bio.*, 49:N155–63, 2004.

[106] S. C. Davis, H. Dehghani, J. Wang, S. Jiang, B. W. Pogue, and K. D. Paulsen. Image-guided diffuse optical fluorescence tomography implemented with laplacian-type regularization. *Opt. Express*, 15:4066–82, 2007.

[107] L. Zhou, A. Jin, and B. Yazici. Fluorescence diffuse optical image reconstruction with a priori information. In *Image Processing (ICIP), 2009 16th IEEE International Conference on*, pages 1741–1744, 7-10 2009.

[108] L. Zhou, B. Yazici, and V. Ntziachristos. Fluorescence molecular tomography reconstruction with a priori anatomical information. In *Proceedings of SPIE Photonics West*, volume 6868O, 2008.

[109] T. Simchony, R. Chellappa, and Z. Lichtenstein. Relaxation algorithms for map estimation of gray-level images with multiplicative noise. *IEEE Transactions on Information Theory*, 36(3):608–13, 1990.

[110] S. Reeves and R. Mersereau. Optimal estimation of the regularization parameter and stabilizing functional for regularized image restoration. *Optical Engineering*, 29(5):446–54, 1990.

[111] N. P. Galatsanos and A. K. Katsaggelos. Methods for choosing the regularization parameter and estimating the noise variance in image restoration and their relation. *IEEE Transactions on Image Processing*, 1(3):322–36, 1992.

[112] P. Charbonnier, L. Blanc-Feraud, G. Aubert, and M. Barlaud. Deterministic edge-preserving regularization in computed imaging. *IEEE Transactions on Image Processing*, 6:298–311.

[113] S. Geman and D. E. McClure. Bayesian image analysis: an application to single photon emission tomography. pages 12–18. Washington, DC, 1985.

[114] T. Hebert and R. Leahy. A generalized em algorithm for 3d bayesian reconstruction from poisson data using gibbs priors. *IEEE Transactions on Medical Imaging*, 8(2):194–202, 1989.

[115] C. Bouman and K. Sauer. A generalized gaussian image model for edge-preserving map estimation. *IEEE Transactions on Image Processing*, 2(3):296–310, 1993.

[116] S. Teboul, L. Blanc-Féraud, G. Aubert, and M. Barlaud. Variational approach for edge-preserving regularization using coupled pde's. *IEEE Transactions on Image Processing*, 7(3):387–97, 1998.

[117] I. I. Gikhman and A. V. Skorokhod. *The theory of stochastic processes I.* Springer-Verlag, Berlin Heidelberg New York, 1974.

[118] Nikolaos Deliolanis, Tobias Lasser, Damon Hyde, Antoine Soubret, Jorge Ripoll, and V. Ntziachristos. Free-space fluorescence molecular tomography utilizing 360° geometry projections. *Opt. Lett.*, 32(4):382–4, 2007.

[119] M. Abramowitz and I. A. Stegun. *Handbook of mathematical functions.* Dover, 1968.

[120] J. D. Jackson. *Classical electrodynamics.* John Wiley and Sons, 1962.

[121] S. Walker, D. A. Boas, and E. Gratton. Photon density waves scattered from cylindrical inhomogeneities: theory and experiments. *Appl. Opt.*, 37(10):1935–1944, 1998.

[122] G. Fubini. Sugli integrali multipli. *Opere scelte*, 2:243–49, 1958.

[123] R. A. Horn. *Matrix analysis.* Cambridge University Press, Cambridge, 1985.

APPENDIX A
LAX-MILGRAM LEMMA

Given a Hilbert space V, a continuous coercive bilinear form $\mathcal{F}(\cdot,\cdot)$ and a continuous linear functional $\mathcal{G}$, there exists a unique $u \in V$ such that

$$\mathcal{F}(v,u) = \mathcal{G}(v), \quad \forall v \in V. \tag{A.1}$$

The proof can be found in [94]. As a direct consequence of this lemma, an upper bound on u can be established as

$$\|u\|_V \leq \frac{1}{Q}\|\mathcal{G}\|_{V^*}, \tag{A.2}$$

where Q is the coercivity constant and V^* is the dual space of V.

APPENDIX B
DEFINITION FOR DUAL NORM OF $\mathcal{F}$

The dual norm of the bilinear form $\mathcal{F}(\cdot,\cdot)$ is defined by [94]

$$\|\mathcal{F}(v,u)\|_{V*} := \sup_{v\neq 0} \frac{|\mathcal{F}(v,u)|}{\|v\|_V}, \tag{B.1}$$

where $\|\cdot\|_{V*}$ denotes the norm of $V^*(\Omega)$ which is the dual space of $V(\Omega)$.

APPENDIX C
COERCIVITY OF $\mathcal{F}$

In the noise-free situation, recall $\mathcal{B} = \mathcal{A}^*\mathcal{A}$, thus $(\mathcal{B}u, u) \geq 0$. From the positive parameters λ_1 and λ_2, we see

$$\begin{aligned} |\mathcal{F}(u,u)| &\geq \lambda_1(u,u) + \lambda_2(\nabla u, \nabla u) \\ &\geq \min(\lambda_1, \lambda_2)\|u\|_1^2 \\ &\geq Q\|u\|_1^2, \end{aligned} \tag{C.1}$$

where $Q := \min(\lambda_1, \lambda_2)$ is the coercivity constant.

In the presence of measurement noise, recall $\mathcal{B}_{LH} = \mathcal{A}^*\Sigma_{\boldsymbol{\varepsilon}}^{-1}\mathcal{A}$, thus $(\mathcal{B}u, u) \geq 0$. For the MAP estimator case, we have

$$\begin{aligned} |\mathcal{F}_{MAP}(u,u)| &\geq (u, \frac{u}{\kappa}) \\ &\geq \frac{\|u\|_0^2}{\|\kappa\|_\infty} \\ &\geq Q\|u\|_0^2, \end{aligned} \tag{C.2}$$

where $Q := 1/\|\kappa\|_\infty$ is the coercivity constant.

For the ML estimator case, we have

$$\begin{aligned} |\mathcal{F}_{ML}(u,u)| &\geq \lambda(u,u) \\ &\geq \lambda\|u\|_0^2 \\ &\geq Q\|u\|_0^2, \end{aligned} \tag{C.3}$$

where $Q := \lambda$ is the coercivity constant.

APPENDIX D
PROOF OF THEOREM 1: UPPER BOUND FOR THE ERROR DUE TO FORWARD PROBLEM DISCRETIZATION

Subtracting $\widetilde{\mathcal{F}}(\psi,\mu)$ from both sides of (2.42) yields:

$$\begin{aligned}\widetilde{\mathcal{F}}(\psi,\widetilde{\mu})-\widetilde{\mathcal{F}}(\psi,\mu) &= \widetilde{\mathcal{G}}(\psi)-\widetilde{\mathcal{F}}(\psi,\mu),\\ \widetilde{\mathcal{F}}(\psi,\widetilde{\mu}-\mu) &= \widetilde{\mathcal{G}}(\psi)-\widetilde{\mathcal{F}}(\psi,\mu). \qquad (D.1)\end{aligned}$$

Adding and subtracting $\mathcal{G}(\psi)$ on the right hand side of (D.1) lead to

$$\begin{aligned}\widetilde{\mathcal{F}}(\psi,\widetilde{\mu}-\mu) &= \widetilde{\mathcal{G}}(\psi)-\widetilde{\mathcal{F}}(\psi,\mu)+\mathcal{G}(\psi)-\mathcal{G}(\psi)\\ &= \widetilde{\mathcal{G}}(\psi)-\widetilde{\mathcal{F}}(\psi,\mu)+\mathcal{F}(\psi,\mu)-\mathcal{G}(\psi)\\ &= (\mathcal{F}-\widetilde{\mathcal{F}})(\psi,\mu)+(\widetilde{\mathcal{G}}-\mathcal{G})(\psi). \qquad (D.2)\end{aligned}$$

Then following the Lax-Milgram Lemma in Appendix A, the error $\mu-\widetilde{\mu}$ is bounded by

$$\begin{aligned}\|\mu-\widetilde{\mu}\|_1 &\leq \frac{1}{\min(\lambda_1,\lambda_2)}\left\|(\mathcal{F}-\widetilde{\mathcal{F}})(\psi,\mu)+(\widetilde{\mathcal{G}}-\mathcal{G})(\psi)\right\|_{1*}\\ &\leq \frac{1}{\min(\lambda_1,\lambda_2)}\left[\left\|(\mathcal{F}-\widetilde{\mathcal{F}})(\psi,\mu)\right\|_{1*}+\left\|(\widetilde{\mathcal{G}}-\mathcal{G})(\psi)\right\|_{1*}\right], \qquad (D.3)\end{aligned}$$

where $\|\cdot\|_{1*}$ is defined in Appendix B.

Clearly,

$$\begin{aligned}(\mathcal{F}-\widetilde{\mathcal{F}})(\psi,\mu) &= (\psi,\mathcal{B}\mu)+\lambda_1(\psi,\mu)+\lambda_2(\nabla\psi,\nabla\mu)-(\psi,\widetilde{\mathcal{B}}\mu)-\lambda_1(\psi,\mu)\\ &\quad -\lambda_2(\nabla\mu,\nabla\psi)\\ &= \left(\psi,(\mathcal{B}-\widetilde{\mathcal{B}})\mu\right). \qquad (D.4)\end{aligned}$$

Note that the dual norm of the bilinear form $(\mathcal{F} - \widetilde{\mathcal{F}})(\psi, \mu)$ is defined by

$$\begin{aligned}
\|(\mathcal{F} - \widetilde{\mathcal{F}})(\psi, \mu)\|_{1*} &:= \sup_{\psi \neq 0} \frac{\left|(\mathcal{F} - \widetilde{\mathcal{F}})(\psi, \mu)\right|}{\|\psi\|_1} \\
&= \sup_{\psi \neq 0} \frac{\left|\left(\psi, (\mathcal{B} - \widetilde{\mathcal{B}})\mu\right)\right|}{\|\psi\|_1} \\
&\leq \sup_{\psi \neq 0} \frac{\|(\mathcal{B} - \widetilde{\mathcal{B}})\mu\|_0 \|\psi\|_0}{\|\psi\|_0}. \qquad \text{(D.5)}
\end{aligned}$$

Hence,

$$\|(\mathcal{F} - \widetilde{\mathcal{F}})(\psi, \mu)\|_{1*} \leq \|(\mathcal{B} - \widetilde{\mathcal{B}})\mu\|_0. \qquad \text{(D.6)}$$

Following [79], we express

$$\begin{aligned}
\|(\mathcal{B} - \widetilde{\mathcal{B}})\mu\|_0 &\approx 2\|\mathcal{A}^*(\mathcal{A} - \widetilde{\mathcal{A}})\mu\|_0 \\
&\approx 2\left\|\sum_{i,j}^{N_S, N_D} g_j^* \phi_i \int_\Omega (g_j^* e_i + \phi_i e_j^*)\mu d\boldsymbol{r}\right\|_0 \\
&\leq 2\max_{i,j} \|g_j^* \phi_i\|_0 \sum_{i,j}^{N_S, N_D} \int_\Omega |(g_j^* e_i + \phi_i e_j^*)\mu| d\boldsymbol{r}, \qquad \text{(D.7)}
\end{aligned}$$

where $e_i = \phi_i - \Phi_i$ and $e_j^* = g_j^* - G_j^*$ are discretization errors in Φ_i and G_j^* respectively. Then if we expand the integral on Ω as a summation of the integrals on the finite element Ω_{mj}, $m = 1, \cdots, N_\Delta^{*j}$ and Ω_{ni}, $n = 1, \cdots, N_\Delta^i$, the upper bound for $\|(\mathcal{B} - \widetilde{\mathcal{B}})\mu\|_0$ can be given as

$$\begin{aligned}
&\|(\mathcal{B} - \widetilde{\mathcal{B}})\mu\|_0 \\
\leq\ & 2\max_{i,j} \|g_j^* \phi_i\|_0 \sum_{i,j}^{N_S, N_D} \left(\sum_{n=1}^{N_\Delta^i} \|g_j^* \mu\|_{0,ni} \|e_i\|_{0,ni} + \sum_{m=1}^{N_\Delta^{*j}} \|\phi_i \mu\|_{0,mj} \|e_j^*\|_{0,mj}\right) \\
\leq\ & 2\max_{i,j} \|g_j^* \phi_i\|_0 \left(\sum_{i=1}^{N_S} \sum_{n,j}^{N_\Delta^i, N_D} \|g_j^* \mu\|_{0,ni} \|e_i\|_{0,ni} + \sum_{j=1}^{N_D} \sum_{m,i}^{N_\Delta^{*j}, N_S} \|\phi_i \mu\|_{0,mj} \|e_j^*\|_{0,mj}\right).
\end{aligned}$$

$$\text{(D.8)}$$

Next, following a similar argument as in (D.5) and (D.6), we obtain

$$\|(\widetilde{\mathcal{G}}-\mathcal{G})(\psi)\|_{1*} \leq \|\tilde{f}-f\|_0. \tag{D.9}$$

Then an upper bound for $\|\tilde{f}-f\|_0$ is given by [79]

$$\begin{aligned}
\|\tilde{f}-f\|_0 &= \left\|(\widetilde{\mathcal{A}}^*-\mathcal{A}^*)\mathbf{\Gamma}\right\|_0 \\
&\approx \left\|\sum_{i,j}^{N_S,N_D}(g_j^* e_i+\phi_i e_j^*)\Gamma_{ij}\right\|_0 \\
&\leq \max_{i,j}|\Gamma_{ij}| \sum_{i,j}^{N_S,N_D}\|g_j^* e_i+\phi_i e_j^*\|_0. \qquad \text{(D.10)}
\end{aligned}$$

A bound for $\max_{i,j}|\Gamma_{ij}|$ can be obtained by using (2.14):

$$\max_{i,j}|\Gamma_{ij}| \leq \max_{i,j}\|g_j^*\phi_i\|_0\|\mu\|_0. \tag{D.11}$$

Finally, to compute an upper bound for $\|\tilde{f}-f\|_0$, we expand the $L^2(\Omega)$ norm computed on Ω as a summation on the finite elements Ω_{mj}, $m=1,\cdots,N_\Delta^j$ and Ω_{ni}, $n=1,\cdots,N_\Delta^i$:

$$\begin{aligned}
&\|\tilde{f}-f\|_0 \\
\leq\ & \max_{i,j}\|g_j^*\phi_i\|_0\|\mu\|_0\left(\sum_{i=1}^{N_S}\sum_{n,j}^{N_\Delta^i,N_D}\|g_j^*\|_{\infty,ni}\|e_i\|_{0,ni}+\sum_{j=1}^{N_D}\sum_{m,i}^{N_\Delta^{*j},N_S}\|\phi_i\|_{\infty,mj}\|e_j^*\|_{0,mj}\right).
\end{aligned} \tag{D.12}$$

In the end, using the discretization error bounds (2.40) and (2.41) leads to the theorem.

APPENDIX E
PROOF OF THEOREM 2: UPPER BOUND FOR THE ERROR DUE TO INVERSE PROBLEM DISCRETIZATION

Taking (2.42) as the starting point, by coercivity we can write

$$\|\widetilde{\mu} - \mu^D\|_1^2 \leq \frac{1}{\min(\lambda_1, \lambda_2)} \widetilde{\mathcal{F}}(\widetilde{\mu} - \mu^D, \widetilde{\mu} - \mu^D).$$

Let $\varphi \in V$, then the above inequality is equivalent to

$$\begin{aligned} \|\widetilde{\mu} - \mu^D\|_1^2 &\leq \frac{1}{\min(\lambda_1, \lambda_2)} \left[\widetilde{\mathcal{F}}(\widetilde{\mu} - \mu^D, \widetilde{\mu} - \varphi) + \widetilde{\mathcal{F}}(\widetilde{\mu} - \mu^D, \varphi - \mu^D) \right] \\ &\leq \frac{1}{\min(\lambda_1, \lambda_2)} \widetilde{\mathcal{F}}(\widetilde{\mu} - \mu^D, \widetilde{\mu} - \varphi), \end{aligned}$$

because $\varphi - \mu^D \in V$ and the error $\widetilde{\mu} - \mu^D$ is orthogonal to the finite-dimensional subspace with respect to the norm induced by the bilinear form [94]. Noting

$$\begin{aligned} \widetilde{\mathcal{F}}(\widetilde{\mu} - \mu^D, \widetilde{\mu} - \varphi) &= \Big(\widetilde{\mu} - \mu^D, \widetilde{\mathcal{B}}(\widetilde{\mu} - \varphi) \Big) + \lambda_1 (\widetilde{\mu} - \mu^D, \widetilde{\mu} - \varphi) \\ &\quad + \lambda_2 \left(\nabla(\widetilde{\mu} - \mu^D), \nabla(\widetilde{\mu} - \varphi) \right), \end{aligned} \tag{E.1}$$

it is clear that

$$\begin{aligned} \|\widetilde{\mu} - \mu^D\|_1^2 &\leq \frac{1}{\min(\lambda_1, \lambda_2)} \Big[\|\widetilde{\mu} - \mu^D\|_0 \|\widetilde{\mathcal{B}}(\widetilde{\mu} - \varphi)\|_0 + \lambda_1 \|\widetilde{\mu} - \mu^D\|_0 \|\widetilde{\mu} - \varphi\|_0 \\ &\quad + \lambda_2 \|\nabla\widetilde{\mu} - \nabla\mu^D\|_0 \|\nabla\widetilde{\mu} - \nabla\varphi\|_0 \Big] \\ &\leq \frac{1}{\min(\lambda_1, \lambda_2)} \Big[\|\widetilde{\mu} - \mu^D\|_1 \|\widetilde{\mathcal{B}}(\widetilde{\mu} - \varphi)\|_0 + \lambda_1 \|\widetilde{\mu} - \mu^D\|_1 \|\widetilde{\mu} - \varphi\|_0 \\ &\quad + \lambda_2 \|\widetilde{\mu} - \mu^D\|_1 \|\nabla\widetilde{\mu} - \nabla\varphi\|_0 \Big]. \end{aligned} \tag{E.2}$$

Canceling $\|\widetilde{\mu} - \mu^D\|_1$ terms on both sides, we have

$$\|\widetilde{\mu} - \mu^D\|_1 \leq \frac{1}{\min(\lambda_1, \lambda_2)} \left[\|\widetilde{\mathcal{B}}(\widetilde{\mu} - \varphi)\|_0 + \lambda_1 \|\widetilde{\mu} - \varphi\|_0 + \lambda_2 \|\nabla\widetilde{\mu} - \nabla\varphi\|_0 \right]. \tag{E.3}$$

Let $\varphi \in V$ be the interpolant of $\widetilde{\mu}$ and $e_\mu = \widetilde{\mu} - \varphi$ be the interpolation error. Then the first term in the bound (E.3) can be expanded as follows:

$$\begin{aligned}
\|\widetilde{\mathcal{B}}(\widetilde{\mu} - \varphi)\|_0 &= \|\widetilde{\mathcal{B}} e_\mu\|_0 \\
&= \left\| \int_\Omega \sum_{i,j}^{N_S,N_D} G_j^*(\cdot)\Phi_i(\cdot)G_j^*(\acute{\boldsymbol{r}})\Phi_i(\acute{\boldsymbol{r}})e_\mu(\acute{\boldsymbol{r}})d\acute{\boldsymbol{r}} \right\|_0 \\
&\leq \sum_{i,j}^{N_S,N_D} \|G_j^*\Phi_i\|_0 \int_\Omega |G_j^*(\acute{\boldsymbol{r}})\Phi_i(\acute{\boldsymbol{r}})e_\mu(\acute{\boldsymbol{r}})| d\acute{\boldsymbol{r}}. \qquad \text{(E.4)}
\end{aligned}$$

Then if we expand the integral on Ω as a summation of the integrals on the finite element Ω_t, $t = 1, \cdots, N$:

$$\begin{aligned}
\|\widetilde{\mathcal{B}}(\widetilde{\mu} - \varphi)\|_0 &\leq \sum_{i,j}^{N_S,N_D} \|G_j^*\Phi_i\|_0 \sum_{t=1}^{N_\Delta} \|G_j^*\Phi_i\|_{0,t} \|e_\mu\|_{0,t} \\
&\leq \max_{i,j} \|G_j^*\Phi_i\|_0 \sum_{i,j}^{N_S,N_D} \sum_{t=1}^{N_\Delta} \|G_j^*\Phi_i\|_{0,t} \|e_\mu\|_{0,t}.
\end{aligned}$$

The remaining two terms in (E.3) can be expressed in a straightforward way:

$$\begin{aligned}
\lambda_1 \|\widetilde{\mu} - \varphi\|_0 &= \lambda_1 \|e_\mu\|_0 \leq \lambda_1 \sum_{t=1}^{N_\Delta} \|e_\mu\|_{0,t}, \\
\lambda_2 \|\nabla\widetilde{\mu} - \nabla\varphi\|_0 &= \lambda_2 \|\nabla e_\mu\|_0 \leq \lambda_2 \sum_{t=1}^{N_\Delta} \|\nabla e_\mu\|_{0,t}.
\end{aligned}$$

Assume that our solution also satisfies $\widetilde{\mu} \in H^2(\Omega)$. Then the bounds for the interpolation error and its gradient on each element can be given by [94]:

$$\begin{aligned}
\|e_\mu\|_{0,t} &\leq C\|\widetilde{\mu}\|_{2,t} h_t^2, \qquad \text{(E.5)} \\
\|\nabla e_\mu\|_{0,t} &\leq C\|\widetilde{\mu}\|_{2,t} h_t, \qquad \text{(E.6)}
\end{aligned}$$

where C is a positive constant, $\|\cdot\|_{0,t}$ and $\|\cdot\|_{2,t}$ are respectively the L^2 and H^2 norms on Ω_t and h_t is the diameter of the smallest ball containing the finite element Ω_t. Finally, substituting (E.5) and (E.6) into (E.3) proves the theorem.

APPENDIX F
SOLUTION OF THE MODEL PROBLEM

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 approximations of ϕ_i and g_j^*. Below, we give the solution in 2D for the forward problem (2.5) [60]:

$$\begin{aligned} g_j(\rho, \rho_s^j; \theta, \theta_s^j) &= \frac{4}{\overline{D}\pi} \left\{ \frac{1}{2} I_0(k_\Omega \rho_<) K_0(k_\Omega \rho_>) \right. \\ &\quad \left. + \sum_{m=1}^{\infty} \cos[m(\theta - \theta_s)] I_m(k_\Omega \rho_<) K_m(k_\Omega \rho_>) \right\}, \end{aligned}$$

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 [119], and $k_\Omega = \sqrt{-K_\Omega^2}$. Under the same conditions, an analytical solution for the adjoint problem (2.9) can be obtained in a similar way. The solution of the problem in three-dimensional can be derived in a similar manner [120, 121].

APPENDIX G
LOWER BOUND ON Θ

Substituting the constraint (3.13) into the bound (2.54), we get

$$\begin{aligned}
\|\widetilde{\mu} - \mu^D\|_1 &\leq \frac{1}{\min(\lambda_1, \lambda_2)} \left(\max_{i,j} \|G_j^* \Phi_i\|_0 \epsilon \sum_{t=1}^{N_\Delta} \sum_{i,j}^{N_S, N_D} \|G_j^* \Phi_i\|_{0,t} h_t \right. \\
&\quad \left. + \lambda_1 \sum_{t=1}^{N_\Delta} \epsilon h_t + \lambda_2 \sum_{t=1}^{N_\Delta} \epsilon \right) \\
&:= \Theta. \qquad \text{(G.1)}
\end{aligned}$$

Next, we compute a lower bound for Θ in (G.1).

$$\begin{aligned}
\Theta &= \frac{1}{\min(\lambda_1, \lambda_2)} \left(\max_{i,j} \|G_j^* \Phi_i\|_0 \epsilon \sum_{t=1}^{N_\Delta} \sum_{i,j}^{N_S, N_D} \|G_j^* \Phi_i\|_{0,t} h_t + \lambda_1 \sum_{t}^{N_\Delta} \epsilon h_t + \lambda_2 \sum_{t}^{N_\Delta} \epsilon \right) \\
&\geq \frac{N_\Delta \epsilon \max_{i,j} \|G_j^* \Phi_i\|_0}{\min(\lambda_1, \lambda_2)} \min_t \sum_{i,j}^{N_S, N_D} \|G_j^* \Phi_i\|_{0,t} h_t + \frac{\lambda_1}{\min(\lambda_1, \lambda_2)} N_\Delta \epsilon \min_t h_t \\
&\quad + \frac{\lambda_2}{\min(\lambda_1, \lambda_2)} N_\Delta \epsilon \\
&\geq N_\Delta \epsilon \left[1 + \min_t h_t + \frac{\max_{i,j} \|G_j^* \Phi_i\|_0}{\min(\lambda_1, \lambda_2)} \min_t \sum_{i,j}^{N_S, N_D} \|G_j^* \Phi_i\|_{0,t} h_t \right]. \qquad \text{(G.2)}
\end{aligned}$$

APPENDIX H
PROOF OF THEOREM 3: UPPER BOUND FOR THE BIAS OF MAP ESTIMATE

Let $\bar{\mu}_{MAP}(\boldsymbol{r}) := \mathrm{E}[\mu_{MAP}(\boldsymbol{r})]$ and $\bar{\mu}^D_{MAP}(\boldsymbol{r}) := \mathrm{E}[\mu^D_{MAP}(\boldsymbol{r})]$, then $\mathrm{E}[e_{MAP}(\boldsymbol{r})] = \bar{\mu}_{MAP}(\boldsymbol{r}) - \bar{\mu}^D_{MAP}(\boldsymbol{r})$. We can further express $\mathrm{Bias}^2[\mu^D_{MAP}]$ in (4.25) as

$$\begin{aligned}\mathrm{Bias}^2[\mu^D_{MAP}] &= \int_\Omega \left|\bar{\mu}_{MAP}(\boldsymbol{r}) - \bar{\mu}^D_{MAP}(\boldsymbol{r})\right|^2 d\boldsymbol{r} \\ &= \|\bar{\mu}_{MAP}(\boldsymbol{r}) - \bar{\mu}^D_{MAP}(\boldsymbol{r})\|_0^2.\end{aligned}$$

It is clear that $\mathrm{Bias}^2[\mu^D_{MAP}]$ is the square of $L^2(\Omega)$ norm of the difference between $\bar{\mu}_{MAP}$ and $\bar{\mu}^D_{MAP}$.

Taking the expectation on both sides of (4.9) and (4.18), and applying Fubini's theorem [122], we can show that $\bar{\mu}_{MAP}$ and $\bar{\mu}^D_{MAP}$ satisfy the following variational forms:

$$\begin{aligned}\mathcal{F}_{MAP}(\psi, \bar{\mu}_{MAP}) &= (\psi, \mathcal{A}^*\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}^{-1}\bar{\boldsymbol{\Gamma}}) + (\psi, \frac{\mu_0}{\kappa}), \quad \forall\psi \in L^2(\Omega), &\text{(H.1)}\\ \widetilde{\mathcal{F}}_{MAP}(\Psi, \bar{\mu}^D_{MAP}) &= (\Psi, \widetilde{\mathcal{A}}^*\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}^{-1}\bar{\boldsymbol{\Gamma}}) + (\Psi, \frac{\mu_0}{\kappa}), \quad \forall\Psi \in V(\Omega), &\text{(H.2)}\end{aligned}$$

where $\bar{\boldsymbol{\Gamma}} = \mathrm{E}[\boldsymbol{\Gamma}]$. Further, we let $\bar{\bar{\mu}}_{MAP} \in L^2(\Omega)$ be the solution of the following approximate inverse problem:

$$\widetilde{\mathcal{F}}_{MAP}(\psi, \bar{\bar{\mu}}_{MAP}) = (\psi, \widetilde{\mathcal{A}}^*\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}^{-1}\bar{\boldsymbol{\Gamma}}) + (\psi, \frac{\mu_0}{\kappa}), \quad \forall\psi \in L^2(\Omega).$$

Then we have

$$\mathrm{Bias}^2[\mu^D_{MAP}] \le \left[\|\bar{\mu}_{MAP} - \bar{\bar{\mu}}_{MAP}\|_0 + \|\bar{\bar{\mu}}_{MAP} - \bar{\mu}^D_{MAP}\|_0\right]^2, \tag{H.3}$$

by the triangular inequality.

For the first term $\|\bar{\mu}_{MAP} - \bar{\bar{\mu}}_{MAP}\|_0$ in (H.3), we follow the similar procedures given in Appendix D and obtain

$$\|\bar{\mu}_{MAP}-\bar{\tilde{\mu}}_{MAP}\|_0 \;\leq\; \|\kappa\|_\infty\left[\left\|(\mathcal{B}_{LH}-\widetilde{\mathcal{B}}_{LH})\bar{\mu}_{MAP}\right\|_0+\left\|(\widetilde{\mathcal{A}}^*-\mathcal{A}^*)\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}^{-1}\bar{\boldsymbol{\Gamma}}\right\|_0\right]. \tag{H.4}$$

For term $\left\|(\mathcal{B}_{LH}-\widetilde{\mathcal{B}}_{LH})\bar{\mu}_{MAP}\right\|_0$, we have [79]

$$\begin{aligned}\left\|(\mathcal{B}_{LH}-\widetilde{\mathcal{B}}_{LH})\bar{\mu}_{MAP}\right\|_0 &\approx 2\left\|\sum_{i,j}^{N_S,N_D}\frac{g_j^*\phi_i}{\sigma^2_{\boldsymbol{\varepsilon},ij}}\int_\Omega(g_j^*e_i+\phi_ie_j^*)\bar{\mu}_{MAP}d\boldsymbol{r}\right\|_0\\ &\leq 2\sum_{i,j}^{N_S,N_D}\frac{\|g_j^*\phi_i\|_0}{\sigma^2_{\boldsymbol{\varepsilon},ij}}\int_\Omega|(g_j^*e_i+\phi_ie_j^*)\bar{\mu}_{MAP}|d\boldsymbol{r},\end{aligned}$$

where $e_i=\phi_i-\Phi_i$ and $e_j^*=g_j^*-G_j^*$. Decomposing the integral on Ω into a summation of the integrals on the finite elements Ω_{mj}, $m=1,\cdots,N_\Delta^{*j}$ and Ω_{ni}, $n=1,\cdots,N_\Delta^i$, which are used to discretize the forward problem, we arrive at

$$\begin{aligned}\|(\mathcal{B}_{LH}-\widetilde{\mathcal{B}}_{LH})\bar{\mu}_{MAP}\|_0 \;\leq\; 2&\left(\sum_{i=1}^{N_S}\sum_{n,j}^{N_\Delta^i,N_D}\frac{\|g_j^*\phi_i\|_0}{\sigma^2_{\boldsymbol{\varepsilon},ij}}\|g_j^*\bar{\mu}_{MAP}\|_{0,ni}\|e_i\|_{0,ni}\right.\\ &\left.+\sum_{j=1}^{N_D}\sum_{m,i}^{N_\Delta^{*j},N_S}\frac{\|g_j^*\phi_i\|_0}{\sigma^2_{\boldsymbol{\varepsilon},ij}}\|\phi_i\bar{\mu}_{MAP}\|_{0,mj}\|e_j^*\|_{0,mj}\right).\end{aligned} \tag{H.5}$$

Similarly, for term $\left\|(\widetilde{\mathcal{A}}^*-\mathcal{A}^*)\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}^{-1}\bar{\boldsymbol{\Gamma}}\right\|_0$, we have its upper bound as

$$\begin{aligned}\left\|(\widetilde{\mathcal{A}}^*-\mathcal{A}^*)\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}^{-1}\bar{\boldsymbol{\Gamma}}\right\|_0 &\approx \left\|\sum_{i,j}^{N_S,N_D}\frac{g_j^*e_i+\phi_ie_j^*}{\sigma^2_{\boldsymbol{\varepsilon},ij}}\bar{\Gamma}_{i,j}\right\|_0\\ &\leq \left(\sum_{i=1}^{N_S}\sum_{n,j}^{N_\Delta^i,N_D}\frac{|\bar{\Gamma}_{i,j}|}{\sigma^2_{\boldsymbol{\varepsilon},ij}}\|g_j^*\|_{\infty,ni}\|e_i\|_{0,ni}\right.\\ &\quad\left.+\sum_{j=1}^{N_D}\sum_{m,i}^{N_\Delta^{*j},N_S}\frac{|\bar{\Gamma}_{i,j}|}{\sigma^2_{\boldsymbol{\varepsilon},ij}}\|\phi_i\|_{\infty,mj}\|e_j^*\|_{0,mj}\right).\end{aligned} \tag{H.6}$$

In the end, substituting (H.5) and (H.6) into (H.4) and using the discretization error bounds (2.40) and (2.41) lead to B_1 and B_2 in Theorem 3.

For the second term $\|\bar{\tilde{\mu}}_{MAP} - \bar{\mu}^D_{MAP}\|_0$ in (H.3), we follow the procedures in Appendix E and obtain

$$\|\bar{\tilde{\mu}}_{MAP} - \bar{\mu}^D_{MAP}\|_0 \;\; \leq \;\; \|\kappa\|_\infty \left[\left\| \widetilde{\mathcal{B}}_{LH}(\bar{\tilde{\mu}}_{MAP} - \varphi) \right\|_0 + \left\| \frac{\bar{\tilde{\mu}}_{MAP} - \varphi}{\kappa} \right\|_0 \right]. \tag{H.7}$$

Let $\varphi \in V(\Omega)$ be the interpolant of $\bar{\tilde{\mu}}_{MAP}$ and $e_\mu := \bar{\tilde{\mu}}_{MAP} - \varphi$ be the interpolation error, we have

$$\begin{aligned} \left\| \widetilde{\mathcal{B}}_{LH}(\bar{\tilde{\mu}}_{MAP} - \varphi) \right\|_0 \;\; &= \;\; \left\| \sum_{i,j}^{N_S,N_D} \frac{G_j^*(\cdot)\Phi_i(\cdot)}{\sigma^2_{\boldsymbol{\mathcal{E}},ij}} \int_\Omega G_j^*(\boldsymbol{\acute{r}})\Phi_i(\boldsymbol{\acute{r}})e_\mu(\boldsymbol{\acute{r}})d\boldsymbol{\acute{r}} \right\|_0 \\ &\leq \;\; \sum_{i,j}^{N_S,N_D} \frac{\|G_j^*\Phi_i\|_0}{\sigma^2_{\boldsymbol{\mathcal{E}},ij}} \int_\Omega |G_j^*(\boldsymbol{\acute{r}})\Phi_i(\boldsymbol{\acute{r}})e_\mu(\boldsymbol{\acute{r}})|d\boldsymbol{\acute{r}} \\ &\leq \;\; \sum_{i,j}^{N_S,N_D} \frac{\|G_j^*\Phi_i\|_0}{\sigma^2_{\boldsymbol{\mathcal{E}},ij}} \sum_{t=1}^{N_\Delta} \|G_j^*\Phi_i\|_{0,t}\|e_\mu\|_{0,t}, \end{aligned} \tag{H.8}$$

and

$$\left\| \frac{\bar{\tilde{\mu}}_{MAP} - \varphi}{\kappa} \right\|_0 \leq \sum_{t=1}^{N_\Delta} \left\| \frac{e_\mu}{\kappa} \right\|_{0,t} \leq \sum_{t=1}^{N_\Delta} \left\| \frac{1}{\kappa} \right\|_{\infty,t} \|e_\mu\|_{0,t}\,. \tag{H.9}$$

Approximating $\bar{\tilde{\mu}}_{MAP}$ by $\bar{\mu}_{MAP}$, and substituting (H.8), (H.9) and the discretization error bound given by

$$\|e_\mu\|_{0,t} \leq C\|\bar{\tilde{\mu}}_{MAP}\|_{1,t}h_t,$$

into (H.7), we obtain B_3 in Theorem 3.

APPENDIX I

PROOF OF THEOREM 4: UPPER BOUND FOR THE VARIANCE OF MAP ESTIMATE

We express Var[μ^D_{MAP}] in (4.26) using μ_{MAP}, μ^D_{MAP}, $\bar{\mu}_{MAP}$ and $\bar{\mu}_{MAP}$ as

$$\begin{aligned}\mathrm{Var}[\mu^D_{MAP}] &:= \int_\Omega \mathrm{E}\left[\left|\mu_{MAP}(\boldsymbol{r}) - \mu^D_{MAP}(\boldsymbol{r}) - [\bar{\mu}_{MAP}(\boldsymbol{r}) - \bar{\mu}^D_{MAP}(\boldsymbol{r})]\right|^2\right] d\boldsymbol{r} \\ &= \int_\Omega \mathrm{E}\left[\left|[\mu_{MAP}(\boldsymbol{r}) - \bar{\mu}_{MAP}(\boldsymbol{r})] - [\mu^D_{MAP}(\boldsymbol{r}) - \bar{\mu}^D_{MAP}(\boldsymbol{r})]\right|^2\right] d\boldsymbol{r}.\end{aligned}$$

Subtracting (H.1) and (H.2) from (4.9) and (4.18), respectively, we obtain

$$\begin{aligned}\mathcal{F}_{MAP}(\psi, \mu_{MAP}(\boldsymbol{r}) - \bar{\mu}_{MAP}) &= (\psi, \mathcal{A}^* \boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}^{-1}(\boldsymbol{\Gamma} - \bar{\boldsymbol{\Gamma}})), \quad \forall \psi \in L^2(\Omega), \qquad \text{(I.1)} \\ \widetilde{\mathcal{F}}_{MAP}(\Psi, \mu^D_{MAP}(\boldsymbol{r}) - \bar{\mu}^D_{MAP}) &= (\Psi, \widetilde{\mathcal{A}}^* \boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}^{-1}(\boldsymbol{\Gamma} - \bar{\boldsymbol{\Gamma}})), \quad \forall \Psi \in V(\Omega). \qquad \text{(I.2)}\end{aligned}$$

Let $\mathbf{e}_{ij} \in \mathbb{R}^{N_S N_D}$, $i = 1, \ldots, N_S$ and $j = 1, \ldots, N_D$, be the unit vector given by

$$\mathbf{e}_{ij} = [0,\, \ldots,\, 1,\, \ldots,\, 0]^T,$$

with only non-zero entry at $[(i-1)N_D + j]$th position. Then we define $\pi_{ij}(\boldsymbol{r}) \in L^2(\Omega)$ and $\Pi_{ij}(\boldsymbol{r}) \in V(\Omega)$ as the solution of the following variational problems:

$$\begin{aligned}\mathcal{F}_{MAP}(\psi, \pi_{ij}) &= (\psi, \mathcal{A}^* \boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}^{-1} \mathbf{e}_{ij}), \quad \forall \psi \in L^2(\Omega), \\ \widetilde{\mathcal{F}}_{MAP}(\Psi, \Pi_{ij}) &= (\Psi, \widetilde{\mathcal{A}}^* \boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}^{-1} \mathbf{e}_{ij}), \quad \forall \Psi \in V(\Omega).\end{aligned}$$

Further, we define $\boldsymbol{\pi}(\boldsymbol{r})$ and $\boldsymbol{\Pi}(\boldsymbol{r})$ as

$$\begin{aligned}\boldsymbol{\pi}(\boldsymbol{r}) &= [\pi(\boldsymbol{r})_{11},\, \ldots,\, \pi(\boldsymbol{r})_{1N_S},\, \pi(\boldsymbol{r})_{21},\, \ldots,\, \pi(\boldsymbol{r})_{N_S N_D}]^T, \\ \boldsymbol{\Pi}(\boldsymbol{r}) &= [\Pi(\boldsymbol{r})_{11},\, \ldots,\, \Pi(\boldsymbol{r})_{1N_S},\, \Pi(\boldsymbol{r})_{21}, \ldots,\, \Pi(\boldsymbol{r})_{N_S N_D}]^T.\end{aligned}$$

Due to the linearity of (I.1) and (I.2), it can be shown that $\mu_{MAP}(\boldsymbol{r}) - \bar{\mu}_{MAP}(\boldsymbol{r})$ and

$\mu^D_{MAP}(\boldsymbol{r}) - \bar{\mu}^D_{MAP}(\boldsymbol{r})$ can be given as

$$\begin{aligned}\mu_{MAP}(\boldsymbol{r}) - \bar{\mu}_{MAP}(\boldsymbol{r}) &= \boldsymbol{\pi}(\boldsymbol{r})^T(\boldsymbol{\Gamma} - \bar{\boldsymbol{\Gamma}}),\\ \mu^D_{MAP}(\boldsymbol{r}) - \bar{\mu}^D_{MAP}(\boldsymbol{r}) &= \boldsymbol{\Pi}(\boldsymbol{r})^T(\boldsymbol{\Gamma} - \bar{\boldsymbol{\Gamma}}).\end{aligned}$$

Then, we have

$$\begin{aligned}\mathrm{Var}[\mu^D_{MAP}] &= \int_\Omega \mathrm{E}\left[\left|[\boldsymbol{\pi}(\boldsymbol{r}) - \boldsymbol{\Pi}(\boldsymbol{r})]^T(\boldsymbol{\Gamma} - \bar{\boldsymbol{\Gamma}})\right|^2\right] d\boldsymbol{r}\\ &= \int_\Omega [\boldsymbol{\pi}(\boldsymbol{r}) - \boldsymbol{\Pi}(\boldsymbol{r})]\, \boldsymbol{\Sigma}_{\boldsymbol{\Gamma}}\, [\boldsymbol{\pi}(\boldsymbol{r}) - \boldsymbol{\Pi}(\boldsymbol{r})]\, d\boldsymbol{r}, \qquad \text{(I.3)}\end{aligned}$$

where $\boldsymbol{\Sigma}_{\boldsymbol{\Gamma}}$ is the covariance matrix of $\boldsymbol{\Gamma}$.

From the property of Gaussian random field, it can be shown that $\boldsymbol{\Gamma}_0$ in our measurement model (4.1) is a multivariate Gaussian random variable statistically independent to the noise $\boldsymbol{\varepsilon}$. More specifically, using the Fubini's theorem, we can derive the mean of $\boldsymbol{\Gamma}_0$ and the covariance between each pair of its entries:

$$\begin{aligned}\mathrm{E}\left[\boldsymbol{\Gamma}_0\right] &= \mathcal{A}\mu_0,\\ \mathrm{Cov}\left[\Gamma_{0,ij}, \Gamma_{0,kl}\right] &= \int_\Omega \kappa(\boldsymbol{r}) a_{ij}(\boldsymbol{r}) a_{kl}(\boldsymbol{r}) d\boldsymbol{r}.\end{aligned}$$

Then, due to the independence between the noise and fluorophore concentration, $\boldsymbol{\Sigma}_{\boldsymbol{\Gamma}}$ can be obtained as the sum of the covariance matrices of $\boldsymbol{\Gamma}_0$ and noise $\boldsymbol{\varepsilon}$. We use $\left(\boldsymbol{\Sigma}_{\boldsymbol{\Gamma}}\right)_{p,q}$, for $p = 1, \ldots, N_S N_D$, and $q = 1, \ldots, N_S N_D$, to denote the entry at the p^{th} row and the q^{th} column of $\boldsymbol{\Sigma}_{\boldsymbol{\Gamma}}$, then we have

$$\left(\boldsymbol{\Sigma}_{\boldsymbol{\Gamma}}\right)_{p,q} = \delta_{pq}\sigma^2_{\boldsymbol{\varepsilon},ij} + \int_\Omega \kappa(\boldsymbol{r}) a_{ij}(\boldsymbol{r}) a_{kl}(\boldsymbol{r}) d\boldsymbol{r},$$

where δ_{pq} is the Kronecker delta function and the indices i, j, k, l, p and q have the following relationship:

$$p = (i-1)N_D + j, \quad q = (k-1)N_D + l,$$

for $i, k = 1, \ldots, N_S$, $j, l = 1, \ldots, N_D$, and $p, q = 1, \ldots, N_S N_D$.

In this respect, we can express (I.3) as

$$\begin{aligned}
&\mathrm{Var}[\mu_{MAP}^{D}] \\
= \;& \int_{\Omega} \sum_{i,j}^{N_S,N_D} \left(\boldsymbol{\Sigma}_{\boldsymbol{\Gamma}}\right)_{(i-1)N_D+j,(i-1)N_D+j} \left|\pi_{ij}(\boldsymbol{r}) - \Pi_{ij}(\boldsymbol{r})\right|^2 d\boldsymbol{r} \\
&+ \int_{\Omega} \sum_{i,j}^{N_S,N_D} \sum_{k,l,kl\neq ij}^{N_S,N_D} \left(\boldsymbol{\Sigma}_{\boldsymbol{\Gamma}}\right)_{(i-1)N_D+j,(k-1)N_D+l} \left[\pi_{ij}(\boldsymbol{r}) - \Pi_{ij}(\boldsymbol{r})\right]^* \left[\pi_{kl}(\boldsymbol{r}) - \Pi_{kl}(\boldsymbol{r})\right] d\boldsymbol{r}.
\end{aligned}$$

Applying Cauchy-Schwarz inequality, we obtain

$$\begin{aligned}
&\mathrm{Var}[\mu_{MAP}^{D}] \\
\leq \;& \sum_{i,j}^{N_S,N_D} \left(\boldsymbol{\Sigma}_{\boldsymbol{\Gamma}}\right)_{(i-1)N_D+j,(i-1)N_D+j} \left\|\pi_{ij}(\boldsymbol{r}) - \Pi_{ij}(\boldsymbol{r})\right\|_0^2 \\
&+ \sum_{i,j}^{N_S,N_D} \sum_{k,l,kl\neq ij}^{N_S,N_D} \left(\boldsymbol{\Sigma}_{\boldsymbol{\Gamma}}\right)_{(i-1)N_D+j,(k-1)N_D+l} \left\|\pi_{ij}(\boldsymbol{r}) - \Pi_{ij}(\boldsymbol{r})\right\|_0 \left\|\pi_{kl}(\boldsymbol{r}) - \Pi_{kl}(\boldsymbol{r})\right\|_0 .
\end{aligned}$$

Since $\boldsymbol{\Sigma}_{\boldsymbol{\Gamma}}$ is positive semi-definite, we have [123]

$$\left|\left(\boldsymbol{\Sigma}_{\boldsymbol{\Gamma}}\right)_{p,q}\right| \leq \left|\left(\boldsymbol{\Sigma}_{\boldsymbol{\Gamma}}\right)_{p,p}\right|^{1/2} \left|\left(\boldsymbol{\Sigma}_{\boldsymbol{\Gamma}}\right)_{q,q}\right|^{1/2},$$

and

$$\begin{aligned}
&\mathrm{Var}[\mu_{MAP}^{D}] \\
\leq \;& \sum_{i,j}^{N_S,N_D} \left(\boldsymbol{\Sigma}_{\boldsymbol{\Gamma}}\right)_{(i-1)N_D+j,(i-1)N_D+j} \left\|\pi_{ij}(\boldsymbol{r}) - \Pi_{ij}(\boldsymbol{r})\right\|_0^2 \\
&+ \sum_{i,j,k,l,ij\neq kl}^{N_S,N_D,N_S,N_D} \left|\left(\boldsymbol{\Sigma}_{\boldsymbol{\Gamma}}\right)_{(i-1)N_D+j,(i-1)N_D+j}\right|^{1/2} \left|\left(\boldsymbol{\Sigma}_{\boldsymbol{\Gamma}}\right)_{(k-1)N_D+l,(k-1)N_D+l}\right|^{1/2} \\
&\left\|\pi_{ij}(\boldsymbol{r}) - \Pi_{ij}(\boldsymbol{r})\right\|_0 \left\|\pi_{kl}(\boldsymbol{r}) - \Pi_{kl}(\boldsymbol{r})\right\|_0 \\
\leq \;& \left[\sum_{i,j}^{N_S,N_D} \left|\left(\boldsymbol{\Sigma}_{\boldsymbol{\Gamma}}\right)_{(i-1)N_D+j,(i-1)N_D+j}\right|^{1/2} \left\|\pi_{ij}(\boldsymbol{r}) - \Pi_{ij}(\boldsymbol{r})\right\|_0\right]^2 . \qquad \text{(I.4)}
\end{aligned}$$

Note that, when the measurements are noise-free and the fluorophore concentration is also deterministic, we have $\boldsymbol{\Sigma}_{\boldsymbol{\Gamma}} = \mathrm{E}\left[(\boldsymbol{\Gamma} - \bar{\boldsymbol{\Gamma}})(\boldsymbol{\Gamma} - \bar{\boldsymbol{\Gamma}})\right] = \mathbf{0}$, thereby $\mathrm{Var}[\mu_{MAP}^{D}] =$

0.

For each $\|\pi_{ij}(\boldsymbol{r}) - \Pi_{ij}(\boldsymbol{r})\|_0$ in (I.4), we can let $\bar{\boldsymbol{\Gamma}} = \mathbf{e}_{ij}$ and follow the same approach in Appendix H to obtain an upper bound for $\|\pi_{ij}(\boldsymbol{r}) - \Pi_{ij}(\boldsymbol{r})\|_0$ given by

$$
\begin{aligned}
& \|\pi_{ij}(\boldsymbol{r}) - \Pi_{ij}(\boldsymbol{r})\|_0 \\
\leq\ & C \cdot \left[\sum_{i'=1}^{N_S} \sum_{n,j'}^{N_\Delta^{i'},N_D} \frac{2\|\kappa\|_\infty \|g_{j'}^* \phi_{i'}\|_0}{\sigma_{\boldsymbol{\varepsilon},i'j'}^2} \|g_{j'}^* \pi_{ij}\|_{0,ni'} \|\phi_{i'}\|_{1,ni'} h_{ni'} \right. \\
& + \frac{\|\kappa\|_\infty}{\sigma_{\boldsymbol{\varepsilon},ij}^2} \sum_{n=1}^{N_\Delta^i} \|g_j^*\|_{\infty,ni} \|\phi_i\|_{1,ni} h_{ni} \\
& + \sum_{j'=1}^{N_D} \sum_{m,i'}^{N_\Delta^{*j'},N_S} \frac{2\|\kappa\|_\infty \|g_{j'}^* \phi_{i'}\|_0}{\sigma_{\boldsymbol{\varepsilon},i'j'}^2} \|\phi_{i'} \pi_{ij}\|_{0,mj'} \|g_{j'}^*\|_{1,mj'} h_{mj'} \\
& + \frac{\|\kappa\|_\infty}{\sigma_{\boldsymbol{\varepsilon},ij}^2} \sum_{m=1}^{N_\Delta^{*j}} \|\phi_i\|_{\infty,mj} \|g_j^*\|_{1,mj} h_{mj} \\
& \left. + \sum_{t=1}^{N_\Delta} \left(\sum_{i',j'}^{N_S,N_D} \frac{\|\kappa\|_\infty \|G_{j'}^* \Phi_{i'}\|_0 \|G_{j'}^* \Phi_{i'}\|_{0,t}}{\sigma_{\boldsymbol{\varepsilon},i'j'}^2} + \|\kappa\|_\infty \left\| \frac{1}{\kappa} \right\|_{\infty,t} \right) \|\pi_{ij}\|_{1,t} h_t \right]. \quad \text{(I.5)}
\end{aligned}
$$

Finally, let $D_{ij} = \left| \left(\boldsymbol{\Sigma}_{\boldsymbol{\Gamma}} \right)_{(i-1)N_D+j,(i-1)N_D+j} \right|^{1/2}$ and substitute (I.5) into (I.4), we arrive at (4.28) in Theorem 4.